ALESSANDRO NICOLI DE MATTOS

COMO MENTIR sutilmente

com gráficos e estatísticas

Com Introdução de
André Luis Ferreira da Silva Bacci

Como Mentir Sutilmente com Gráficos e Estatísticas

(do original Informação é Prata, Compreensão é Ouro)

Um guia para o alfabetismo gráfico e estatístico

por
Alessandro Nicoli de Mattos

1ª Edição - Abr/2015

Este livro está disponível gratuitamente em:
http://archive.org/details/ComoMentirSutilmente

Este livro também pode ser baixado gratuitamente em diferentes formatos de eBook em:
https://www.smashwords.com/books/view/536031

Também está disponível gratuitamente no iBooks, Livraria Cultura, Barnes&Nobles e outros distribuidores.

Sites:
Página no Facebook: **http://www.facebook.com/419386594853051/**
Página no Google+: **http://plus.google.com/118410405608511378703/**

Onde baixar o livro Informação é Prata, Compreensão é Ouro:
http://www.smashwords.com/books/view/281255

Versão em PDF com formatação de livro impresso em:
http://www.archive.org/details/InformacaoEPrataCompreensoEOuro

Capa: André Luis Ferreira da Silva Bacci e Alessandro Nicoli de Mattos

Contato com o autor: comomentirsutilmente@gmail.com

Gênero: Não-Ficção

Palavras chaves: estatísticas, gráficos

LICENÇA DE USO E OUTRAS INFORMAÇÕES

Este trabalho está disponibilizado como *emailware*. Isto significa que não há custo para o leitor desde que o trabalho seja utilizado nas condições descritas abaixo na Licença *Creative Commons*. O único pedido que o *emailware* faz é que o leitor que gostou deste trabalho envie um e-mail para o autor com suas sugestões, comentários, críticas ou elogios. É uma forma rápida e barata de demonstrar apreço pelo trabalho de outras pessoas que enriquecem a Internet com informações úteis e gratuitas.

comomentirsutilmente@gmail.com

Sumário

Introdução

É difícil de fazer a introdução de um livro o qual você admira... Afinal uma introdução visa adiantar a um novo leitor o que ele vai encontrar a frente, um *mapa*, porém um mapa descrito por alguém que já trilhou o caminho. Como lidar com os *spoilers*?

Não ajuda o fato da admiração do livro recair sobre sobre o conjunto da obra. Pois é assim a minha admiração ao extenso livro **A Informação é Prata, o Conhecimento é Ouro** do Alessandro Nicoli de Mattos. É o tipo de obra subversiva, que altera profundamente o seu modo de pensar *sem que você se dê conta*. É um truque sutil, mas nem por isso pouco eficiente. Como um *meme* particularmente virótico ou um *inception* perfeitamente implantado, a obra se incorpora à mente do leitor, numa simbiose perfeita, e por isso mesmo, invisível.

Mas efetiva como ela é, o livro **A Informação é Prata, o Conhecimento é Ouro** é realmente extenso, tratando de assuntos bem variados. Nesse mundo de informação rápida e de super especialização isso acaba sendo uma desvantagem, um contrassenso: o livro que justamente ajudaria a lidar com a frenética Era da Informação não passa ele mesmo pelo critério de leitura dos consumidores que mais necessitam de seus conhecimentos.

Eu percebi isso durante cursos e palestras onde os convivas, munidos de internet e da "segunda tela", rapidamente consultavam e entravam em contato com a obra frequentemente citada, só para em seguida se verem desanimados a enfrentar um livro de mais de 100.000 palavras.

E ante a esse desânimo seguia-se sempre a recomendação que fosse lido pelo menos o capítulo 10 da obra.

Eis que apesar a admirar o conjunto da obra, esse capítulo é o que mais proximidade tem com os assuntos tratados em minhas palestras e cursos, a saber, mercado financeiro e investimentos. Ao tratar mais especificamente do lado bom (e não tão bom assim) da representação da informação em gráficos e informações estatísticas, esses capítulos caem como uma luva no contexto de informações financeiras resumidas em gráficos modernos, em todas suas variadamente criativas (ou alucinadas) formas.

O autor, talvez já reconhecendo essa questão, tinha desde antes destacado parte do livro num blog, magnificamente intitulado **Como Mentir Sutilmente**. Reuniu aí os *maus* exemplos em design citados no livro maior.

Funcionava para acessar exemplos específicos, mas parecia faltar algo... Por um lado, tinha a ausência dos bons exemplos, que servissem de contra peso. Por outro lado os *posts* estanques

pareciam diminuir a capacidade comunicativa. Em outras palavras, sem contexto ou linearidade a mensagem ficava diluída.

* * *

Foi pensando em tudo isso que surgiu a ideia de fazer um meio termo. Aproveitar o ótimo título do blog num livro menor, concentrado na questão de informações visuais, tanto pelo lado positivo como também o lado sinistro da questão. Uma obra coesa o suficiente para ser mais eficaz e curta o suficiente para não ser iniciada com desânimo. Em vez de filtrar o livro maior já entregar um livro menor, focado e atualizado.

Assim surge **Como Mentir Sutilmente**, o livro, esta obra que lês agora.

O foco continua recaindo sobre o uso desleixado ou deliberadamente maligno da representação visual de informações puras ou estatísticas (e assim fazendo jus ao nome). Infelizmente é mais fácil citar os maus exemplos que os bons, até por uma questão de disponibilidade.

Mas em cada exemplo segue-se uma discussão detalhada do que os tornam melhores ou piores. Pois o objetivo não é simplesmente elogiar ou tripudiar dos casos colecionados, mas antes *discutir* em como eles são elogiáveis ou deploráveis, respectivamente contribuindo ou dificultando o entendimento da informação subjacente.

Esse é o grande mote desta obra menor. Melhorar a habilidade do leitor na interpretação e compreensão de informações dispostas em gráficos e outras representações artísticas proporcionais. Tão simples ou intuitiva como possa parecer a questão, ela na verdade engloba várias ciências inteiras: há muito mais num simples gráfico que supõe nossos ingênuos instintos.

Semelhante situação ocorre no contexto de informações estatísticas, onde questões de coleta, parcialidade e comunicação tornam-se tão ou mais importante que os dados informados: é preciso um senso crítico desenvolvido para diferenciar "estatística" de "propaganda nua e crua". Opostos como podem ser em teoria, essas duas formas acabam unidas numa terrível quimera no marketing moderno.

E depois de ler essa obra você estará muito mais capacitado em lidar com a aparente (mas enganosa) simplicidade das informações escondidas nessas formas resumidas.

André Luis Ferreira da Silva Bacci

Dedicatória

Este livro é dedicado a todas as pessoas que já se sentiram enganadas por gráficos e estatísticas.

Também dedico este livro ao meu filho **Leonardo** e à minha esposa **Laura**, que são tão importantes em minha vida.

Agradecimentos

Agradeço ao **André Bacci** pelo incentivo em realizar este livro, e por todo o seu trabalho em torná-lo realidade. É sempre bom encontrar mentes com as quais podemos trabalhar com prazer.

Sobre Estatísticas, Gráficos e Mentirosos

Estatísticas e gráficos são usados constantemente para representar valores numéricos. No entanto, a maioria das pessoas crê que eles estão fora de sua capacidade de compreensão ou tem uma imagem negativa deles. A aquisição de uma "mentalidade estatística" é suficiente para eliminar muito desses problemas.

Estamos cercados de dados estatísticos e gráficos; ligando a TV, abrindo as revistas ou jornais, navegando na Internet, nos *spams* nos nossos *e-mails*, no trabalho ou quando vamos às compras. Eles estão sempre lá, tentando provar um ponto de vista, tentando nos vender algum produto ou ideia ou tentando nos convencer a fazer algo.

Mas ao contrário do que muitos pensam, os números podem ser usados para enganar tanto quanto as palavras, e até ainda mais facilmente, pois os números e gráficos carregam consigo uma aura de científico ou indecifrável. Muitas pessoas desistem de entendê-los antes mesmo de tentar, por julgarem não terem os conhecimentos necessários para tal. Os números atemorizam.

E o quadro parece pior quando descobrimos que mais pessoas têm problemas em interpretar os números do que as palavras; há um índice elevado de **"anumerismo"**. Os dados do INAF[01] mostram que apenas 23% da população brasileira possuem aptidões plenas em matemática, e que apenas 26% têm aptidões plenas em letramento. A diferença pode ser pequena, mas de qualquer forma indica níveis muito baixos no Brasil de capacidade plena de entender números e palavras. E este é um problema em todo o mundo, mesmo em países desenvolvidos.

Quem sofre de algum grau de "anumerismo" deixa de aproveitar a facilidade que os números oferecem, por ser uma das formas mais fáceis de assimilar informação. Os números com seus valores absolutos são capazes de resumir aspectos relevantes da realidade, muitos dos quais afetam diretamente a nossa vida.

Além disso, entender os números e suas estatísticas e gráficos, nos ajuda a compreender que eles não são portadores de uma verdade divina e indiscutível. A partir do momento que começamos a entendê-los e questioná-los, superamos a imagem de superioridade incompreensível que carregam, e somos tomados por uma sensação de posse desses números, e passamos a nos sentir à vontade com eles. Tão à vontade, que percebemos que a maioria dos números que nos cercam podem ser jogados fora, pois não são tão importantes quanto parecem ou quanto fazem parecer.

Quantos números a mídia e outras formas de comunicação jogam sobre nós todos os dias, e, no entanto, não temos a menor ideia de como isso nos afeta? Quantas vezes recebemos estes números com a reverência de monges, pois eles parecem ser de suma importância, mas não fazemos a

mínima ideia do que significam, ou não conseguimos abstrair a sua ordem de grandeza? Preocupamo-nos em saber o aumento da temperatura nos polos, o índice das bolsas de valores ou o superavit do trimestre, mas a maioria de nós não sabe ou não teria interesse em saber a relação entre estes números e os impactos ambientais ao nosso redor, o desempenho da economia ou se um 1 bilhão é pouco ou muito como superavit na nossa economia. Quantas vezes a sua vida foi afetada pela pressão barométrica, o resultado da balança comercial ou o aumento na população de baleias?

"Quando você pode medir o que você está falando e expressar isso em números, você sabe alguma coisa sobre isso; mas quando você não consegue medi-lo, quando você não consegue expressá-lo em números, seu conhecimento é do tipo escasso e insatisfatório". - Lord Kelvin, físico e matemático britânico (1824-1907)

Os números só ganham significado quando podem ser relacionados a algo que pode ser captado visceralmente. Se isso não ocorrer, os números tornam-se inúteis. Isto é especialmente verdade para números grandes.

Números excessivamente grandes, como a distância entre o Sol e a Terra, ou excessivamente pequenos, como o tamanho de uma bactéria, são difíceis de compreender, pois não temos com que correlacioná-los.

Se você tivesse que adivinhar a idade da **Terra** e a idade do **Universo**, em quais números pensaria? **Milhares** de anos? **Milhões** de anos? **Bilhões** de anos? Todas essas opções são igualmente abstratas para a maioria das pessoas, mesmo que entre estes números haja diferenças de milhares de vezes. Se você fizer essa pergunta para muitas pessoas, irá receber respostas que podem variar de 10 mil anos a 10 bilhões de anos. Estes números são tão grandes, e estes períodos de tempo são tão abstratos e gigantes comparados com o tempo de vida de uma pessoa e com os períodos que estamos acostumados a lidar quando estudamos história ou lembramo-nos de eventos passados, que não conseguimos estabelecer uma relação entre o tempo que conhecemos e a idade da Terra ou do Universo.

Quer saber qual a idade da Terra e do Universo? A idade da Terra é estimada em 4,5 bilhões de anos, e a idade do Universo é estimada em 13,7 bilhões de anos.

Estatísticas na mídia, na publicidade e em estudos

Estatísticas são usadas extensivamente na mídia, na publicidade e em todos os tipos de estudo. Não somente porque os números são capazes de informar de uma maneira que as palavras não conseguem, ou sumarizar uma grande massa de dados em uma só afirmação, mas também porque eles podem ser facilmente manipulados e expressos da maneira que for mais conveniente,

defendendo qualquer ponto de vista, e, contudo, sem estar tecnicamente mentindo. **Você não pode provar qualquer fato com estatísticas, mas pode sempre encontrar algo bom ou mal para dizer sobre o assunto.**

Estamos atolados por números e estatísticas que nos atingem por todos os lados, afirmando verdades e gritando por uma importância que não questionamos, por acreditar que se eles estão lá, devem estar corretos e devem ser importantes para nós. São tantas estatísticas querendo provar todo tipo de fato, que muitas até mesmo se contradizem.

"Quando você mata um é uma tragédia, quando você mata dez milhões é uma estatística". - Joseph Stalin, estadista soviético (1878-1953)

Raramente questionamos estes números por considerarmos que se eles estão lá, são corretos e válidos, ou pelo menos alguém já teve o trabalho de verificá-los. Mas isto muitas vezes não é verdade; não passa de informação falsa ou simplesmente não informação. Números que são publicados sem qualquer referência ou verificação podem induzir as pessoas a ações e comportamentos prejudiciais.

Mesmo números que sempre acreditamos serem de suma importância, como os índices das bolsas de valores, como o **IBOVESPA**, por exemplo, são só a média de um conjunto de ações de grandes empresas e não representam o mercado de ações como um todo.

A publicidade é um campo em que sempre temos que questionar o interesse da fonte da informação sobre o que está sendo afirmado. Mesmo nos casos em que estatísticas e avaliações são creditadas a laboratórios independentes, possivelmente os números certos foram escolhidos com um propósito ou algumas informações foram omitidas para transmitir a mensagem certa.

Não é difícil se deparar com comerciais em revistas, jornais, televisão, rádio, Internet ou até em *e-mails spams* patrocinados por fontes suspeitas e com afirmações do tipo (todas elas são apenas exemplos fantasiosos, mas relembram bastante as publicidades que estão por aí):

"Aveia ajuda a baixar as taxas de colesterol em 35% dos casos", em um estudo de um laboratório "independente" patrocinado por uma indústria de cereais.

"Chocolates ajudam a evitar cáries", uma descoberta creditada a uma marca de chocolates.

"Tomar café diariamente ajuda a melhorar a memória em 22%", em pesquisa de uma associação de produtores de café.

"Tomar café diariamente aumenta em 67% as chances de ter gastrite", em pesquisa de uma

associação de produtores de leite.

"Tomar vinho tinto todos os dias aumenta em 5 anos a expectativa de vida", associação de vinícolas.

"Esse novo espremedor de laranja tira 28% a mais de suco que os outros", na caixa do eletrodoméstico sendo vendido no supermercado.

"Comer peixe 4 vezes por semana diminui em 58% a chance de um ataque cardíaco", associação dos comerciantes de peixe.

"Leite de cabra do Himalaia aumenta o nível de potência sexual masculina em 86% depois de 3 meses de uso diário", importadores de leite de cabra do Himalaia.

Os mais desatentos podem não perceber, mas em todas essas afirmações aparentemente claras, sucintas e diretas ao ponto, estão sérios problemas de manipulação de números e omissão de dados importantes, além é claro, do óbvio interesse do produtor na divulgação dessas informações.

"Nunca se mente tanto como antes das eleições, durante uma guerra e depois de uma caçada". - Otto von Bismarck, estadista alemão (1815-1898)

O interesse do produtor geralmente está em exagerar alguma nuance de seus dados estatísticos para corroborar a sua opinião ou ponto de vista. É o que mais comumente acontece nos meios noticiosos, que tendem a exagerar os quadros que as estatísticas sugerem para tornar as matérias e reportagens mais interessantes e consequentemente mais vendáveis. E é difícil acreditar que essas distorções são motivadas por pura ignorância ou preguiça, e que não há uma dose de culpa do produtor, pois quase sempre as situações são exageradas, e quase nunca minimizadas.

Esses produtores de informação esperam que você engula qualquer ponto de vista baseado em números escolhidos a dedo e em estatísticas infundadas. Como Huff[02] explica: *"Parece que o repórter passou pra frente algumas palavras sem se questionar o que elas significam, e espera-se que você as leia tão sem crítica quanto eles para a feliz ilusão que elas te dão de ter aprendido algo. Isto tudo é remanescente de uma velha definição de método de lecionamento de instrução na sala de aula: o processo pelo qual o conteúdo de um livro de texto do instrutor é transferido para o caderno de notas do estudante sem passar através das cabeças de nenhum dos dois"*. Mas não é assim com quem tem uma mentalidade estatística.

Mentalidade estatística

Compreender as aplicações da análise estatística é um dos conhecimentos mais úteis na Era da Compreensão. São inúmeras as ocasiões em que esta habilidade pode salvar o consumidor de

informações de armadilhas, ou evitar que o produtor engane o consumidor por pura ignorância ou falta de atenção. E também é uma habilidade necessária para compreender o mundo a nossa volta e exercer plenamente a nossa cidadania.

*"Pensamento estatístico será um dia tão necessário para uma cidadania eficiente como a habilidade de ler e escrever".*H. G. Wells, escritor britânico (1866-1946)

Quando eu era criança li que a **expectativa de vida média** a partir do momento do nascimento durante o Império Romano era de 28 anos; logo, imaginei que as pessoas morriam por volta dessa idade, e me perguntava como era a vida de pessoas que sabiam que iam morrer tão jovens. Li também que a expectativa de vida média no mundo no final do século XIX era de 37 anos e que a atual é por volta de 67 anos. Como as pessoas passaram a viver em média 30 anos a mais em apenas um século? Será que houve mudanças genéticas em nós ao longo da história recente que nos fizeram viver mais? Essas dúvidas me pareciam bem incoerentes.

"Um conhecimento de estatísticas é como o conhecimento de línguas estrangeiras ou de álgebra; ela pode se provar útil a qualquer hora e sob qualquer circunstância". - Arthur Lyon Bowley, economista e estatístico britânico (1869-1957)

Não, não houve mudanças genéticas recentes que nos fizeram ter vidas mais longas. As pessoas na época do Império Romano eram capazes de viver tanto tempo quanto as pessoas de hoje; é só conferir a idade a que chegaram poetas, filósofos, artistas ou políticos romanos ou gregos que morreram de causas naturais.

A partir do momento que compreendi que essas estatísticas de expectativa de vida não refletiam uma média da longevidade do ser humano, e sim a média do tempo de vida de uma determinada população, as dúvidas desapareceram.

A evolução dos números nas estatísticas de expectativa de vida deve-se à forma como ela é calculada. Ela considera qualquer tipo de morte, por acidente, assassinato, doenças, guerras, e o mais importante, as mortes dos recém-nascidos e crianças, cujas taxas são altas em sociedades pouco desenvolvidas, como as sociedades antigas. Conforme o mundo foi se desenvolvendo, curas para doenças foram criadas e as condições de vida melhoraram, resultando, entre outros, numa queda acentuada da mortalidade infantil, e assim aumentando a expectativa de vida média quando considerada a partir do momento do nascimento; no entanto as pessoas que conseguem sobreviver às causas de morte não naturais continuam morrendo com idade avançada, como há 2.000 anos.

Depois de toda essa confusão e de descobrir como esta estatística é calculada e o que ela realmente significa, pude adquirir uma mentalidade estatística sobre este assunto. Agora sei que se em um país

havia dois habitantes, e um morreu com 99 anos, e o outro com apenas 1 ano de idade, a expectativa de vida média desse país será a simples média aritmética das idades de morte, ou seja, de 50 anos, embora esta média esteja longe das idades reais de óbito. Essa média tem uma grande diferença para os valores de suas parcelas, e por isso pode-se dizer que é uma estatística com dispersão alta.

Em algumas sociedades pouco desenvolvidas, para evitar que as altas taxas de mortalidade infantil distorçam essas estatísticas, a média de vida é calculada a partir dos 5 anos de idade, ao invés do momento do nascimento. E isso me parece óbvio e correto agora, que entendo como a estatística de vida é calculada. Da próxima vez que eu ver uma média da expectativa de vida, vou questionar se ela é a partir do nascimento ou a partir dos 5 anos de idade.

Nós nos deixamos influenciar por informações numéricas e não as questionamos, o que nos impede de compreendê-las realmente. Então o mais provável é que sejamos enganados. Assim, é melhor ignorar uma informação que não se entende, do que tentar agir baseado nela.

A maioria das pessoas não estudou muito estatística e nem têm grande facilidade para matemática mais complexa. No entanto, têm a impressão errada de que são necessários muitos conhecimentos avançados de matemática para entender análises estatísticas. As pessoas nem chegam a considerar entender estatística, por considerá-la acima de suas capacidades de compreensão.

Mas para compreender a maioria das aplicações da estatística, principalmente na mídia, um conhecimento básico é suficiente. Qualquer pessoa pode facilmente adquirir uma **mentalidade estatística**, compreendendo o *significado* de alguns termos, como média, mediana, moda, desvio-padrão, curva normal, probabilidades, erro padrão da média, erro provável, amplitude total, correlação, universo de amostragem, erros de amostragem, limite de segurança entre outros. Com isso poderá compreender os resultados das estatísticas, mesmo que nunca tenha aprendido a executar qualquer um dos cálculos de estatística.

"Existem três tipos de mentiras: mentiras, mentiras abomináveis e estatísticas". - Bejamin Disraeli, político e escritor britânico (1804-1881)

No livro **How to Lie With Statistics**[03] (Como mentir com estatísticas), **Darrel Huff** explica conceitos básicos de estatística em um texto leve, curto e agradável de ler, com ilustrações engraçadas. Apresenta muitos exemplos de situações reais em que as estatísticas foram usadas para enganar, confundir, simplesmente foram mal feitos ou sofreram de erros comuns, sempre levando a conclusões enganosas ou imprecisas.

Escrito por um jornalista autônomo, que não tinha nenhuma especialização em estatística, este livro é um exemplo de como se pode adquirir uma "mentalidade estatística", ao mesmo tempo em que

não se precisa recorrer a mais do que a álgebra básica e a conceitos fáceis de compreender.

Não é à toa (ou por uma ironia, visto que o autor não é um estatístico) que este livro é considerado o texto sobre estatística mais lido no mundo. Escrito em 1954, mas com temas ainda perfeitamente atuais, o livro já vendeu mais de 1,5 milhões de cópias na sua edição em inglês.

Nele, Huff[04] afirma: *"A linguagem secreta das estatísticas, tão atraente em uma cultura voltada a fatos, é empregada para sensacionalizar, inflar, confundir e simplificar em excesso. Métodos estatísticos e termos estatísticos são necessários para relatar a massa de dados das tendências sociais e econômicas, condições de negócios, pesquisas de 'opinião' e o censo. Mas sem escritores que usem as palavras com honestidade e entendimento e leitores que saibam o que elas significam, o resultado só pode ser um absurdo semântico".*

"Primeiro consiga os seus fatos; então os distorça ao seu bel prazer". - Mark Twain, escritor americano (1835-1910)

Mas muitas pessoas não se interessam por estudar estatística, por terem dela um conhecimento intuitivo e que acreditam suficiente. Mas saber como as estatísticas são calculadas nos permite entender o que seus resultados realmente significam e nos permite experimentar uma sensação de posse do que está sendo informado.

Entenda mais as estatísticas

Ao estudar um grande volume de dados, muitas vezes as únicas conclusões valiosas são aquelas fornecidas pela estatística, por seus parâmetros como o desvio-padrão, média, mediana ou moda, por exemplo. Quando se deve consumir uma grande massa de dados, a tarefa é muito difícil sem um parâmetro sumarizador.

Tome como exemplo o conjunto de dados abaixo. Eles são um conjunto de números aleatórios, entre 0 e 1, bastante adequados para este exemplo por sua simplicidade, mas que poderiam ser substituídos por qualquer outro conjunto de números.

Olhando para este conjunto de números não conseguimos tirar nenhuma conclusão; eles parecem somente um emaranhado de números com três casas decimais, sem significado.

0,321	0,883	0,680	0,709	0,287	0,414	0,121
0,351	0,118	0,481	0,184	0,891	0,140	0,188
0,162	0,019	0,744	0,198	0,482	0,883	0,312
0,874	0,662	0,996	0,099	0,569	0,195	0,554
0,290	0,196	0,635	0,449	0,243	0,009	0,739
0,044	0,380	0,140	0,468	0,817	0,597	0,461
0,189	0,639	0,902	0,847	0,289	0,477	0,087
0,793	0,225	0,657	0,823	0,120	0,924	0,140
0,685	0,769	0,089	0,538	0,240	0,843	0,597

No entanto, podemos reduzir este conjunto de dados a alguns parâmetros estatísticos, que nos fornecem informações sumarizantes:

O **número mais alto é 0,996**. O **número mais baixo é 0,009**. Agora sabemos a faixa que estes dados abrangem.

A **média aritmética**, que é a soma de todos os números divididos pela quantidade de números, é **0,458**. A sua **mediana**, que é o número central, ou seja, apresenta a mesma quantidade de números com valores menores e valores maiores, é **0,461**. Note que essas duas médias têm valores muito parecidos, típico das distribuições naturais, como a altura de pessoas, as notas das provas em uma sala de aula ou os níveis de colesterol em determinada população, e também para as distribuições aleatórias, como essa.

A **moda é 0,140**, que é o número que ocorre com mais frequência, aparecendo 3 vezes.

O **percentil de 25% é 0,192**, ou seja, o conjunto de 25% dos números com valores mais baixos está abaixo desse valor. O **percentil de 75% é 0,697**, ou seja, os 75% dos números com valores mais baixos estão abaixo desse valor.

O **desvio-padrão é ±0,288**. Ele é uma medida da dispersão dos dados em torno da média aritmética, quando a distribuição segue a curva no formato de um sino, típica das distribuições naturais. Ele é apresentado como um valor positivo e negativo (±) em torno da média. Esta medida significa que 33% dos números da amostra estão dentro de uma faixa de 0,288 abaixo da média, e outros 33% estão dentro de uma faixa de 0,288 acima da média.

Então podemos sumarizar o conjunto de dados acima nos seguintes parâmetros estatísticos:

Valor máximo: 0,996
Valor mínimo: 0,009
Média aritmética: 0,458
Mediana: 0,461
Moda: 0,140
Percentil de 25%: 0,192

Percentil de 75%: 0,697
Desvio-padrão: ±0,288

Estes parâmetros nos fornecem uma boa síntese do conjunto de dados acima, da qual podemos tirar conclusões; é claro que não têm a mesma precisão e nem substituem o conjunto de dados, mas fornecem uma boa dica do que ele é.

E quanto mais parâmetros estatísticos usarmos, melhor e mais clara será a representação daquele conjunto de dados. Se tivéssemos apresentado somente a média aritmética, o que é bem comum nas estatísticas por aí, teríamos uma visão muito mais deficiente do conjunto de dados. À medida que adicionamos novos parâmetros, que podem ser muitos outros além dos que foram apresentados acima, podemos formar uma visão mais clara do todo.

Mas não se engane. Os parâmetros estatísticos podem esconder erros graves. As médias, relações, tendências e gráficos não são sempre o que parecem; podem esconder alguma informação ou mostrar o que não existe.

Um portador comum de erros e mentiras é a **amostragem**. A amostragem vai determinar a qualidade dos resultados estatísticos. Não são raros os casos de estudos ou pesquisas que tiram conclusões baseadas em amostras pequenas, que fornecem resultados pouco sólidos. Quanto maior a amostragem, menores são os erros estatísticos. Huff[05] afirma que *"o procedimento de amostragem é o coração de grande parte das estatísticas que você encontra sobre todo tipo de assunto"*.

Mas os truques mais desonestos com a amostragem são aqueles que selecionam os dados mais convenientes para os cálculos dos parâmetros estatísticos, o que é agravado pelas facilidades oferecidas pelo computador. Agora é fácil escolher entre dezenas de combinações de dados, gráficos e testes estatísticos, e selecionar para a publicação aqueles que melhor defendem o ponto de vista do produtor. Esta seleção de dados é conhecida como "mineração dos dados" e é parecida com as técnicas dos mágicos, que escolhem o que revelar e o que esconder para manter a atenção do iludido público. O consumidor deve sempre considerar que provavelmente o que está sendo mostrado são os dados minerados, ou a combinação de dados que melhor serve aos propósitos do produtor.

E **não se deixe enganar pela terminologia**: muitas vezes índices e estatísticas têm o mesmo nome, mas isso não significa que são equivalentes, e nem que podem ser colocadas lado a lado e serem comparados. Muitas vezes as terminologias vêm de contextos totalmente diferentes e apresentam critérios e cálculos distintos. É o caso de estatísticas realizadas por órgãos governamentais, que

levam em consideração o contexto de seus próprios países, e geralmente não podem ser diretamente comparados com os mesmos índices de outros países.

Como exemplo, o índice de desemprego pesquisado e calculado pelo governo da Alemanha para seu país pode ser incomparável ao índice de desemprego estimado pelo governo brasileiro; são países com contextos totalmente diferentes, e os critérios para considerar uma pessoa desempregada podem ser diferentes, como com relação ao tempo sem emprego necessário para ser considerada desempregada, a consideração sobre os trabalhos autônomos, outras fontes de renda que não de um emprego, situação de emprego do cônjuge, se está procurando trabalho ou não, entre outros. Para o governo dos EUA, por exemplo, pessoas sem emprego, mas que estão ativamente procurando por trabalho, não são consideradas desempregadas.

Outro problema comum é o uso de **termos vagos**, que podem se aplicar a mais de um parâmetro estatístico, e que podem ser escolhidos de acordo com a conveniência do produtor. Sempre questione a definição das medidas que são usadas naquela estatística em particular.

Um exemplo de termo vago muito usado é várias vezes citado por Huff[06] em seu livro. Ele descreve o uso enganoso que se pode fazer da palavra **"média"**: *"Um truque comum é usar a palavra 'média' para diferentes tipos de média, pois este termo é bem vago. Este é um truque comumente usado, algumas vezes inocentemente, mas geralmente intencionalmente, por pessoas querendo influenciar a opinião pública ou vender espaços de publicidade. Quando te dizem que alguma coisa é uma média você não sabe muito sobre isso até que possa descobrir quais dos tipos comuns de média é – média, mediana ou moda".*

A figura abaixo, preparada por Huff, é uma ótima ilustração das diferenças entre média, mediana e moda. Neste caso, é a análise estatística do salário dos trabalhadores em uma fábrica.

*Figura extraída de **A Produção de Informações Estratégicas**, de **Platt**, que foi adaptada do livro **How to Lie with Statistics**, de **Huff**. Uma ótima lustração das diferenças entre média, mediana e moda.*[07]

Note que neste caso, as diferenças entre os 3 tipos de média são bem acentuadas, e cada média pode se prestar a defender um tipo diferente de argumento. Os sindicalistas podem usar a mediana ou a moda como "média" para defender que a maioria dos trabalhadores ganha pouco e merece um aumento. Enquanto que os patrões podem usar a média aritmética, que é a mais alta devido à influência dos altos salários de uns poucos funcionários da liderança, para defender que a "média" de salários na fábrica é alta e que reajustes salariais não são necessários.

Usar somente a "média" como um argumento estatístico pode ser um indício de estatística tendenciosa, principalmente se a média não for especificada. Huff[08] escreve: *"Geralmente uma*

média – seja ela uma média, mediana, especificada ou não – é de tal simplificação excessiva que é pior do que inútil. Saber nada sobre o assunto é frequentemente mais saudável que saber o que não é, e um pequeno aprendizado pode ser uma condição perigosa". É como na história do estudante que somou todas as pessoas do mundo e descobriu que em média cada pessoa tem um seio e um testículo, embora ninguém se encaixe nesta descrição; esta é uma média que não tem utilidade.

E também há o **erro estatístico**, que é uma estimativa do erro entre o que foi observado e o valor que se espera confirmar. É como nas pesquisas de intenção de voto nas eleições, em que é informado que o erro da pesquisa pode levar os resultados 2% ou 3% para cima ou para baixo. Este valor é importante quando comparamos valores semelhantes, pois, se considerado o erro estatístico, as conclusões sobre o quadro geral podem ser diferentes, como o resultado esperado de uma eleição ou a comparação dos resultados de um teste. Lembre-se que comparações entre valores com diferenças pequenas só têm significado se você mantiver em mente este mais ou menos, mesmo (ou especialmente) quando isso não é informado.

Desconfie de estatísticas com **resultados muito precisos**. Acredita-se que 93,2151734% de todas as estatísticas alegam uma precisão de resultados que não pode ser justificada pelo método empregado. Na afirmação anterior parece óbvio que tal precisão não passa de uma piada. Mas num momento de desatenção ou de leitura sem crítica, não é difícil creditarmos como verdadeiras as estatísticas que "estimam" que 40,87% das intenções de voto são para determinado candidato, que uma pasta de dente é 23,19% mais eficiente em evitar cáries que as concorrentes ou que um energético aumenta o nível de energia em 32,82%. Como eles podem saber disso, com essa precisão? Quais métodos foram empregados? Estes métodos justificam a divulgação de números com tal resolução? Tais resoluções geralmente advêm de cálculos matemáticos, usados para calcular médias ou outro parâmetro qualquer, mas não provam e nem indicam a precisão do método estatístico utilizado. O problema é que números arredondados parecem falsos, e até certo grau são mesmo, mas na maioria dos casos são mais que suficiente para informar, com o devido grau de precisão que a estatística oferece, e são mais fáceis de lembrar e comparar. Mas números muito precisos geralmente impressionam as pessoas, e por isso são muito usados.

Outro tipo de informação cuja ausência pode levar a desentendimentos é aquela que diz **a faixa ou amplitude da amostra ou o seu desvio da média**, este último geralmente representado pelo desvio-padrão.

As estatísticas na maioria das vezes se concentram nas médias, em torno de um ponto central. O interesse na maioria dos casos é com o que ocorre no corpo principal dos resultados, no que é frequente, normal e mediano. No entanto, algumas vezes são os extremos que nos interessam; é

aquilo que se afasta ao máximo do ponto central, tanto em uma das extremidades quanto em outra. É o que ocorre com quem quer determinar o peso máximo que uma ponte deve suportar, ou identificar o ponto fraco em um sistema.

Como no exemplo da minha falha em compreender as estatísticas de expectativa de vida, se fosse informado as idades máximas, eu não teria incorrido no mesmo erro. Um monte de confusão pode ser evitado se à média for acrescentada uma indicação da faixa de variação.

Questione as estatísticas

Nem todas as estatísticas que nos são apresentadas podem ser analisadas e certificadas como dignas de credibilidade, como uma substância é analisada e certificada num laboratório de química. Huff[09] sugere 5 perguntas que se forem feitas, podem ser facilmente respondidas e fornecem uma dica do que estamos vendo. Com isso podemos evitar aprender informações que na verdade não são bem assim.

1) Quem diz isso?

Huff[10] explica: *"O primeiro item a procurar é a tendência – o laboratório com alguma coisa a provar por causa de uma teoria, da reputação ou de um pagamento; o jornal cujo objetivo é uma boa história; trabalhadores e patrões com os níveis de salário em jogo"*. Lembre-se que mesmo os laboratórios independentes podem ter interesses em jogo.

Geralmente essas estatísticas tendenciosas são caracterizadas por afirmações enganosas, ambíguas ou que não podem ser provadas; por seleção de dados favoráveis e supressão de desfavoráveis; unidades de medida faltando ou que mudam sem aviso; seleção de uma referência errada para comparação levando a um resultado que interesse mais; uso de parâmetros não especificados, como no caso da palavra "média" que também é usada para cobrir os parâmetros de "mediana" e "moda" e é usada de acordo com o interesse do produtor.

2) Como ele sabe?

Questione como foi possível o produtor levantar estes dados. Muitas vezes os dados são de amostras tendenciosas, com indivíduos selecionados para que o resultado seja o esperado pelo produtor, ou que a amostragem é tão pequena que qualquer resultado poderia ser possível, mas somente aquele resultado que interessa ao produtor é publicado.

Pergunte-se: É a amostragem grande o suficiente para permitir qualquer conclusão confiável? É grande o suficiente para ter um significado? Também desconfie de números com precisões injustificadas, que geralmente tendem a enganar simulando uma autoridade sem fundamento.

Ou ainda medições de valores difíceis de quantificar, como o nível de energia de uma pessoa, o desempenho sexual ou a sensação de conforto. Para estas medições de aspectos humanos existem metodologias que garantem os melhores resultados possíveis; verifique se foi usada uma metodologia adequada, ou se o produtor baseou-se apenas na entrevista de algumas pessoas. E lembre-se que neste tipo de pesquisa o entrevistado também pode distorcer o resultado. São poucas as pessoas que se sentem à vontade em admitir que têm disfunções sexuais, problemas psicológicos ou até mau hálito.

3) O que está faltando?

O que está faltando é difícil identificar, mas fornece uma boa dica de como a estatística está sendo usada para confundir ou de uma falha em seus argumentos. A ausência de algum parâmetro ou valor, particularmente se a fonte está interessada no resultado, é suficiente para jogar suspeita no argumento inteiro.

Procure por uma correlação fornecida sem uma medida de confiabilidade (erro provável, erro padrão), e será um indício de uma estatística não muito séria. Também procure por uma média de variedade não especificada, em qualquer assunto onde a média, a mediana e a moda podem diferir substancialmente.

4) Alguém mudou o assunto?

Huff[11] em seu livro chama a atenção: "Quando avaliando uma estatística, procure por uma mudança em algum lugar entre os valores brutos e a conclusão. Um assunto é muitas vezes informado como outro".

Geralmente isso ocorre pela suposição de uma correlação que não pode ser provada. A mudança de assunto ocorre quando se apresenta que algo "é devido a" algum motivo, formando uma conclusão sobre os dados. Mas essa relação é somente assumida como sendo verdadeira, e não pode ser provada.

É como no exemplo em que os dados estatísticos mostram que o número de casos registrados de uma determinada doença aumentou em relação à pesquisa anterior. Mas isso não necessariamente significa que a doença está se espalhando. Pode ter havido uma mudança nos critérios para registro dessa doença que aumentam o número de casos que são registrados, ou o aumento do conhecimento sobre a doença permite a identificação de casos quando antes não se sabia a causa; ou o sistema de saúde está cobrindo uma área maior com mais pessoas. Tudo isso pode explicar o aumento de casos registrados, sem significar, necessariamente, que a doença está se espalhando. Isso invalida a conclusão que a correlação entre o aumento de casos registrados e a disseminação da doença seja

verdade. Os dados estatísticos mostram um aumento dos casos registrados, e só; ela não mostra um aumento nos casos da doença, embora possa sugerir isso; a correlação ficou por conta da imaginação, desatenção ou desonestidade do produtor da informação.

A mudança de assunto ocorre quando presumimos que algo significa automaticamente outro. **Correlação não implica causação, mas é somente um indício que ela pode existir**.

Essa correlação "forçada" é especialmente estimulada quando temos estatísticas que não se ajustam completamente bem ao propósito que queremos, então correlacionamos o que queremos saber com os dados que temos. Como no exemplo anterior, a estatística dos casos registrados da doença foi utilizada para estimar o número de casos da doença, podendo levar a uma conclusão errônea sobre o alastramento de uma epidemia, mesmo que os números da estatística estejam corretos. No entanto, muitas vezes esses são os melhores números que podemos conseguir, e o produtor tem que julgar por si mesmo se eles são realmente apropriados para o fim desejado, e, se for o caso, informar adequadamente o consumidor sobre as premissas utilizadas para a correlação.

5) Isso faz sentido?

Uma pergunta simples, que requer somente um pouco de exercício da imaginação para ser respondida, e, no entanto, pode evitar conclusões gravemente erradas.

A pergunta *"isso faz sentido?"* provoca uma análise objetiva que muitas vezes irá derrubar uma estatística baseada em suposições erradas, pelo simples motivo de elas não apresentarem uma coerência básica ou fugirem do bom senso, o que geralmente pode ser notado desde o princípio. Estatísticas com erros desse tipo só conseguem sucesso quando a aura mágica dos números causa uma suspensão do bom senso. Exemplos são as correlações absurdas, amostras claramente tendenciosas e fortes interesses do produtor.

Outro exemplo de absurdo é a extrapolação incontrolada. Quando fazemos previsões baseadas em tendências do passado, ignoramos que para isso ocorrer está implícito que tudo o mais deverá estar igual e que a tendência tem que continuar a ocorrer. Mas no mundo real, por algum motivo, o ambiente que nos cerca se recusa a permanecer igual, o que, aliás, se não ocorresse faria a vida ser muito chata.

Um aumento grande e constante no passado não significa necessariamente que o aumento continuará a ocorrer na mesma velocidade, pois o mercado fica saturado com produtos, as pessoas chegam ao seu limite de consumo em determinada área, o crescimento das famílias chega a um limite, o crescimento populacional esbarra em limitações naturais e de infraestrutura, assim como a economia, entre outros.

Uso honesto de gráficos

Os gráficos são poderosos mapas que nos guiam através dos números. Eles oferecem um quadro visual dos números que antes pareciam intimidadores, mas agora parecem muito mais agradáveis e convidativos. Gráficos mostram visualmente quantidades medidas por meio do uso combinado de pontos, linhas, formas geométricas, algum sistema coordenado, números, símbolos, palavras e cores.

Na pequena extensão de um olhar temos todas as informações que os números trazem desenhadas na forma de linhas, áreas, pontos e muitos outros recursos gráficos. Podemos facilmente compará-los, perceber tendências ou as relações entre as parcelas e o todo, e podemos realizar estas tarefas com muito mais rapidez.

O uso de figuras abstratas para mostrar números é uma invenção recente, com suas primeiras aparições significativas ocorrendo por volta do final do século XVIII, principalmente em estudos científicos e trabalhos de engenharia e estatística. E isto ocorreu bem depois de triunfos da matemática como os logaritmos, coordenadas Cartesianas, cálculo integral e o básico da teoria da probabilidade. Talvez tenha sido assim devido à diversidade de habilidades requeridas para desenhar gráficos – conhecimentos matemáticos, pensamento estatístico, habilidades artísticas - numa época em que não existia o auxílio de computadores, calculadoras e ferramentas gráficas.

Mas gráficos não são simples substituições de conjuntos de números; quando usados apropriadamente, tornam-se instrumentos para o raciocínio sobre informações quantitativas. Por isso, gráficos não são simplesmente veículos neutros de comunicação. Eles não somente mostram dados. Embora algumas pessoas possam defender a integridade dos gráficos com o ditado de que *"os números não mentem"*, essa pequena peça da sabedoria popular não é aplicável aos gráficos, embora possa ser verdade para os números.

Gráficos são maneiras de reduzir e interpretar dados para acompanhar uma mensagem. Existe pelo menos uma maneira de interpretar os dados em um gráfico, mesmo que não seja o modo que o desenhista escolheu. A conclusão disso, como Tufte[12] defende, é que **há maneiras corretas e erradas de mostrar dados; há visualizações que revelam a verdade e há visualizações que não**.

Então não é por acaso que gráficos são extensivamente usados na mídia, na publicidade, em relatórios técnicos, relatórios anuais de empresas e em qualquer outro lugar onde seja necessária uma representação sumarizada e visual de uma grande quantidade de números, e, ao mesmo tempo, o produtor está interessado em transmitir uma mensagem específica.

Gráficos e suas relações com as pessoas

Os gráficos também são um meio efetivo de contornar as dificuldades que as pessoas têm em lidar com números. Ao transformar números em representações visuais, a informação parece mais acessível e menos ameaçadora. Huff[13] analisa a importância de desenhar gráficos para evitar os problemas que as pessoas têm com os números: *"Talvez soframos de um trauma induzido pela aritmética do primário. Qualquer que seja a causa, isto cria um problema real para o escritor que anseia por ser lido, o publicitário que espera que sua propaganda venda os produtos, o editor que quer que seus livros e revistas sejam populares. Quando números em forma tabular são um tabu e palavras não irão fazer o trabalho bem, o que é frequentemente o caso, só existe uma reposta sobrando: desenhe uma figura"*.

Mas não são todos que conseguem tirar proveito desses mapas que nos guiam para a compreensão. O INAF[14], um estudo sobre o alfabetismo funcional no Brasil, mostrou que somente os que tinham alfabetização plena em matemática é que conseguiam demonstrar familiaridade com representações gráficas como mapas, tabelas e gráficos, e menos da metade dos entrevistados declarou prestar atenção nos gráficos que acompanham matérias de jornais ou revistas. E como acontece com muitos assuntos que não conseguimos compreender, ou enxergamos com uma aura de científico, acabamos por aceitá-los sem questionar.

Os gráficos também sofrem de outro mal. Além de parecerem incompreensíveis para muitas pessoas (ou por causa disso), eles carregam consigo uma imagem de enganosos e manipuladores, e, portanto, pouco confiáveis e que não merecem a nossa atenção. Mas Tufte[15] alerta que os gráficos não são o único meio que sofre desse mal: *"Para muita gente, a primeira palavra que vem à mente quando se pensa em gráficos estatísticos é 'mentira'. Não há dúvida de que alguns gráficos realmente distorcem os dados essenciais, tornando difícil para o observador ver a verdade. Mas, nesse ponto, os gráficos não diferem das palavras, pois qualquer meio de comunicação pode ser usado para iludir. Os gráficos não são especialmente vulneráveis; na realidade, quase todos nós temos ótimos detectores de mentiras gráficas que nos ajudam a ver através das fraudes"*.

Talvez essa imagem ruim tenha surgido com o uso que os governos absolutistas e ditatoriais constantemente fazem das informações divulgadas pelos meios de comunicação de seus países, não por acaso estatais; eles formam a opinião do povo por meio de informações enganosas, manipuladas e com dados importantes omitidos, que logo caem na incredibilidade dos consumidores da informação.

Outra explicação para isso é a antiga suposição de que gráficos só serviam para mostrar o óbvio aos ignorantes. Foi somente no final da década de 1960 que os gráficos começaram a ser usados como ferramentas para raciocinar sobre informações quantitativas. Esta elevação de prestígio contribuiu

para popularizá-los na mídia e no cotidiano das pessoas, para alegria dos que têm inclinações matemáticas, e para desespero dos indivíduos verbais que insistem que não conseguem compreender números.

No entanto, não se deve perder de vista que gráficos não são nada mais que mapas, no sentido de que tornam padrões e tendências perceptíveis e compreensíveis. Por isso, o princípio diretor da produção de gráficos deve ser o cuidado com o conteúdo, e não com o estilo.

Esse princípio diretor que aponta para o conteúdo, em lugar do estilo, procura eliminar o que Tufte chama de "lixo de gráfico" (*"chartjunk"*), que de forma simples é qualquer traço de tinta colocado no papel que não tenha a função de informar, e, por conseguinte, irá inevitavelmente atrapalhar a compreensão do conteúdo. Geralmente são enfeites desnecessários colocados para embelezar o gráfico e satisfazer a ânsia estética do produtor.

Princípios da elegância gráfica

Além da orientação sobre o lixo de gráfico, Tufte[16] sugere cinco princípios que podem guiar o produtor a criar gráficos melhores. São os princípios da excelência gráfica:

- A excelência gráfica é a apresentação bem desenhada de dados interessantes – uma questão de substância, de estatística e de desenho;
- A excelência gráfica consiste de ideias complexas comunicadas com clareza, precisão e eficiência;
- A excelência gráfica é aquela que dá ao observador o maior número de ideias no menor tempo com o mínimo de tinta no menor espaço;
- A excelência gráfica é quase sempre multivariada;
- A excelência gráfica requer contar a verdade sobre os dados.

Esses princípios guiam o produtor para longe de armadilhas gráficas, que distorcem e obscurecem os gráficos e o entendimento do consumidor. Muitas dessas armadilhas serão mostradas mais à frente.

Os gráficos mais comuns

As quatro formas mais comuns de gráfico usadas para comunicar estatísticas são gráficos de barras, circulares, de evolução e tabelas. Cada um deles se presta a tipos específicos de informação.

A escolha de determinada forma deve ter por base o tipo de informação a ser mapeada. Se a escolha for equivocada, ou se a confecção dos gráficos não for cuidadosa, estes quatro simples e comuns tipos de gráfico podem gerar muito confusão.

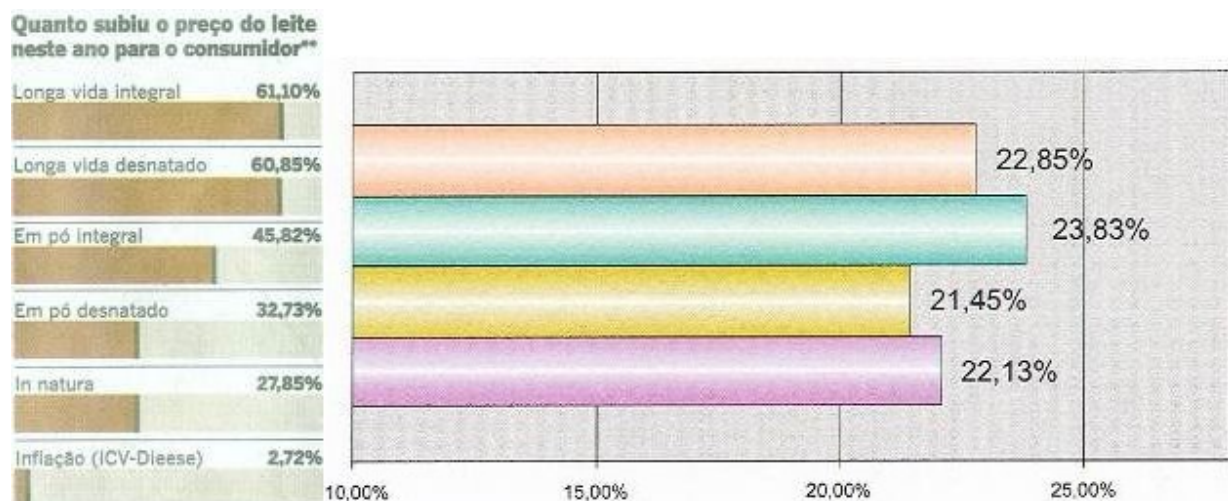
É claro que existem dezenas de tipos de representações gráficas, algumas desenvolvidas para aplicações específicas, como índices econômicos e de bolsas de valores, estudos estatísticos, engenharia, entre outros. Mas estes quatro tipos de gráficos correspondem à maior parte das

representações sendo publicadas por aí, em especial na mídia e na publicidade.

Gráficos de barras

Os gráficos de barras são usados para comparar quantidades. A altura ou comprimento das barras representa a quantidade. Se você tem uma reta, pode facilmente verificar se ela é duas vezes maior do que outra colocada ao lado.

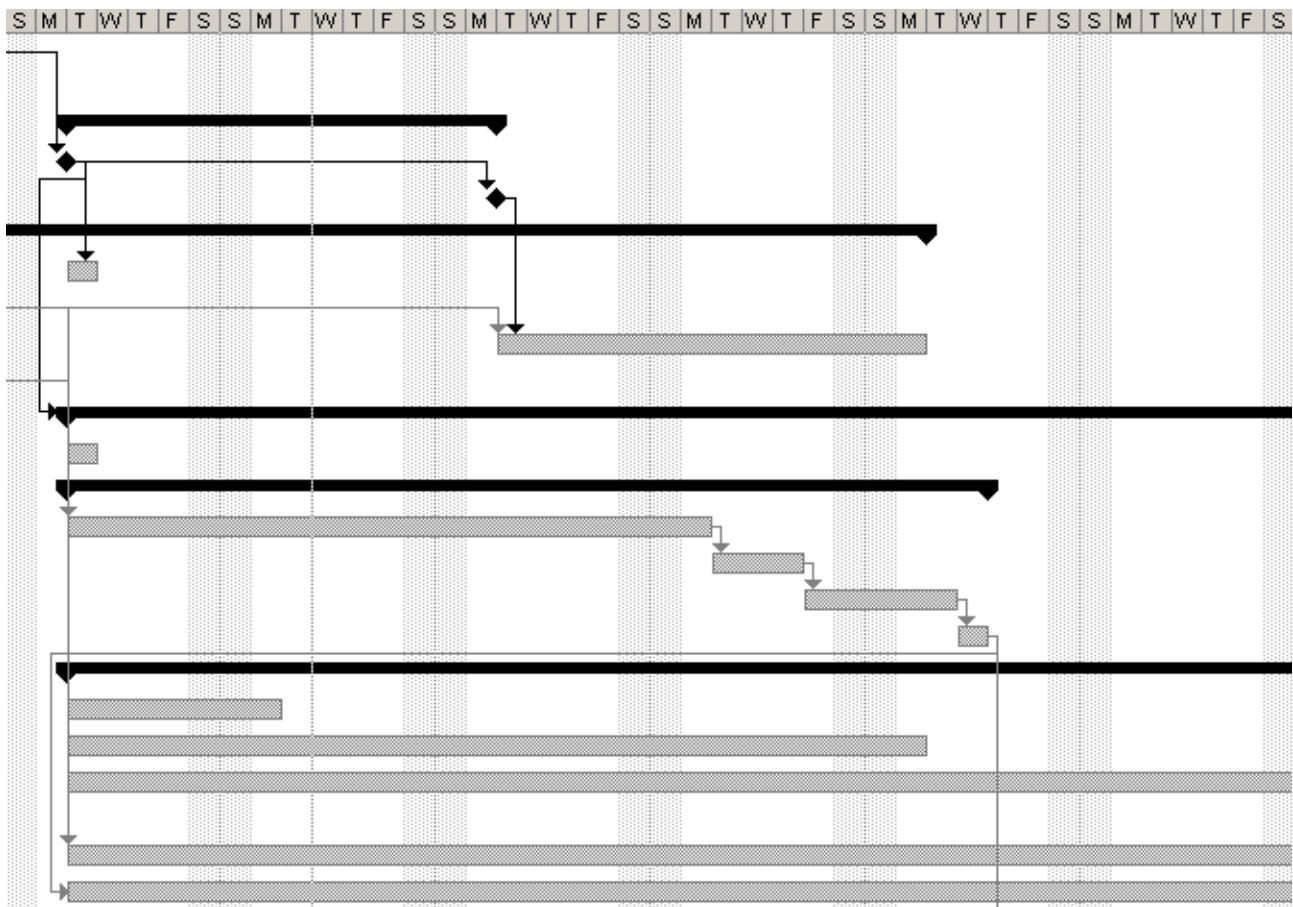
Os gráficos de barras podem apresentar tanto barras verticais quanto horizontais, como nos exemplos abaixo.

Exemplo de um bom gráfico de barras vertical, mostrando a taxa de investimento no Brasil em relação ao PIB. Note que embora as variações sejam pequenas, o produtor do gráfico fez a escolha correta de manter a base das barras em zero.[17]

Exemplos de gráficos de barras horizontais. À esquerda,[18]*, mostrando o aumento de preço de diversos tipos de leite em relação a um índice que mede a inflação. À direita,*[19] *a rentabilidade de diversos tipos de carteiras de investimento; a base não zero faz parecer que as diferenças entre os valores são maiores do que realmente são.*

No entanto, enquanto barras horizontais são um meio matematicamente correto de representar quantidades, eles não comunicam a ideia eficazmente, pois desafiam a noção de que esquerda e

direita denotam a passagem do tempo, e não quantidades, que são mais bem representadas com variações verticais.

Como Jones[20] afirma, *"deitar um gráfico vertical no seu lado pode ser uma forma sutil de mentira porque você confunde o significado intuitivo de em cima e em baixo e de esquerda e direita"*. E complementa: *"O formato horizontal é enganador por que a audiência irá precisar de um momento – mesmo que breve – para se ajustar ao desenho contra intuitivo"*.

Mas as barras horizontais são um excelente meio de mostrar durações de tempo. Um exemplo dessa aplicação das barras horizontais são os gráficos de **Gantt**, que são largamente usados para gerenciar projetos, e geralmente podem ser gerados por programas de gerenciamento de projetos, como o **Microsoft Project**. Eles mostram a duração de cada tarefa e a relação no tempo entre cada uma; barras sobrepostas mostrando o tempo planejado para cada tarefa e o tempo realmente usado indicam facilmente a eficiência da execução da tarefa; além disso, o efeito do deslocamento de uma ou mais tarefas que são interligadas a outras pode ser claramente percebido por barras que avançam e empurram as demais à frente, mostrando o efeito do deslocamento dessas tarefas em particular com o prazo final para conclusão do projeto.

Exemplo de gráfico de gerado pelo ***Microsoft Project****. O comprimento das barras indica a duração da tarefa, e as setas indicam as suas inter-relações.*

Mas os gráficos de barras não se restringem ao uso de barras paralelas para comparação de comprimento. Uma aplicação bem interessante é apresentada em **Understanding USA**, no qual o arquiteto de informação **Nigel Holmes** apresenta um conceito novo em gráficos de barras. O valor é representado pela largura das barras, e várias barras podem se fundir para representar uma união ou se dividir em várias outras para representar uma divisão, com as suas novas larguras representando o resultado dessas operações. Este tipo de representação é chamado de **cosmográfico**, e é particularmente eficiente em mostrar como entradas e saídas iguais são distribuídas, como por exemplo, em assuntos relacionados a dinheiro.

Este tipo de gráfico está sendo usado em relatórios anuais de algumas empresas, para representar as receitas e os gastos, pois é uma ótima maneira de conseguir visualmente relacionar as parcelas (diversas receitas) e o todo (o total de receitas e as reservas), e depois o todo e suas novas parcelas (os gastos e investimentos que saíram das receitas e reservas).

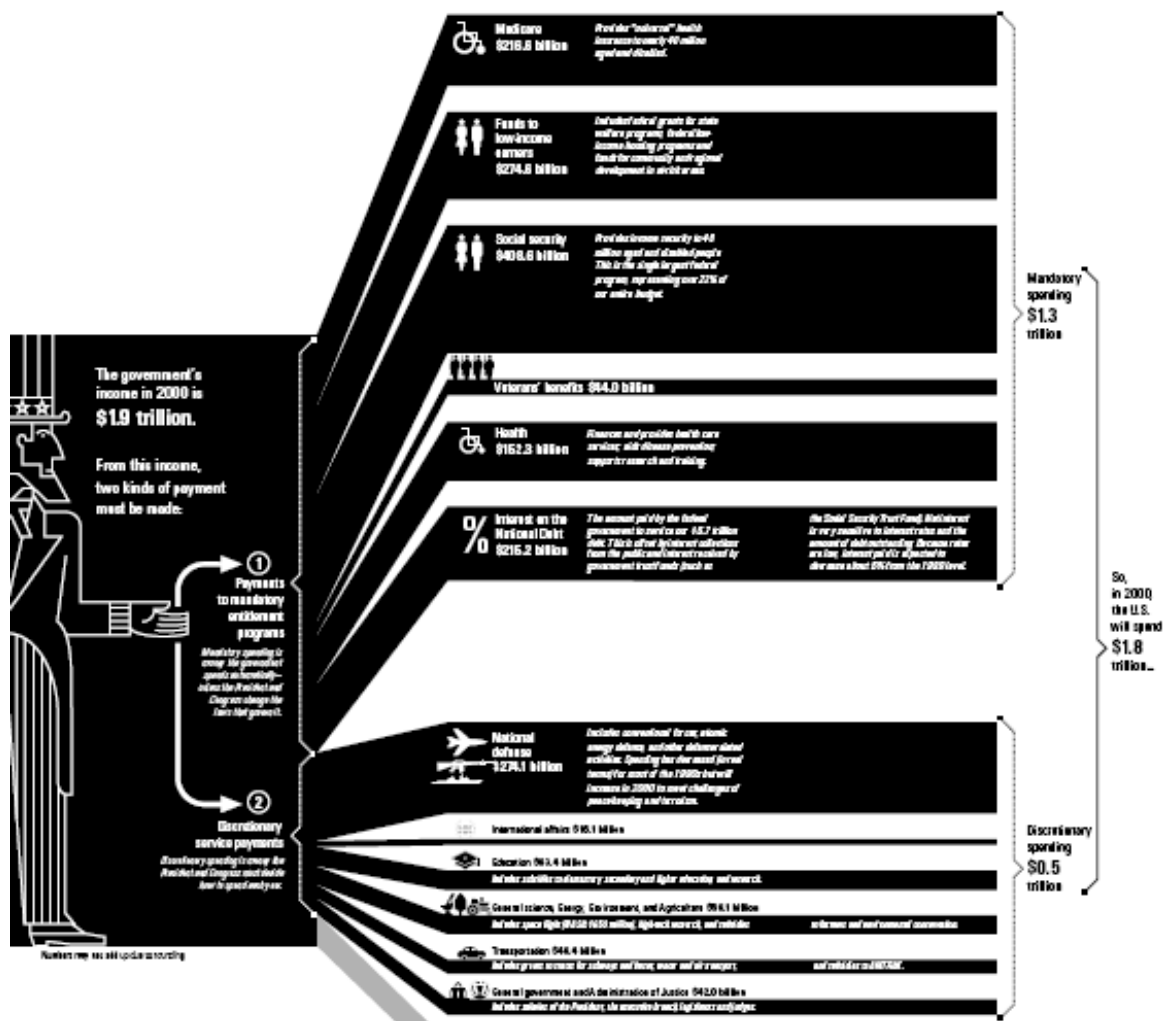



Gráfico de barras de ***Understanding USA***[21]*, de* ***Nigel Holmes****, mostrando as despesas federais a partir da arrecadação. A espessura das barras denota o valor numérico, que pode ser visualmente comparado com a origem e as outras parcelas.*

Mas nem todos os gráficos de barras são honestos. A comparação das barras confia na nossa

percepção visual, e esta pode ser facilmente enganada. Olhe com suspeita para barras que mudam suas larguras junto com o seu comprimento quando representando somente um fator, para as barras que são truncadas ou nas quais elas desenham objetos tridimensionais cujos volumes não são fáceis de comparar.

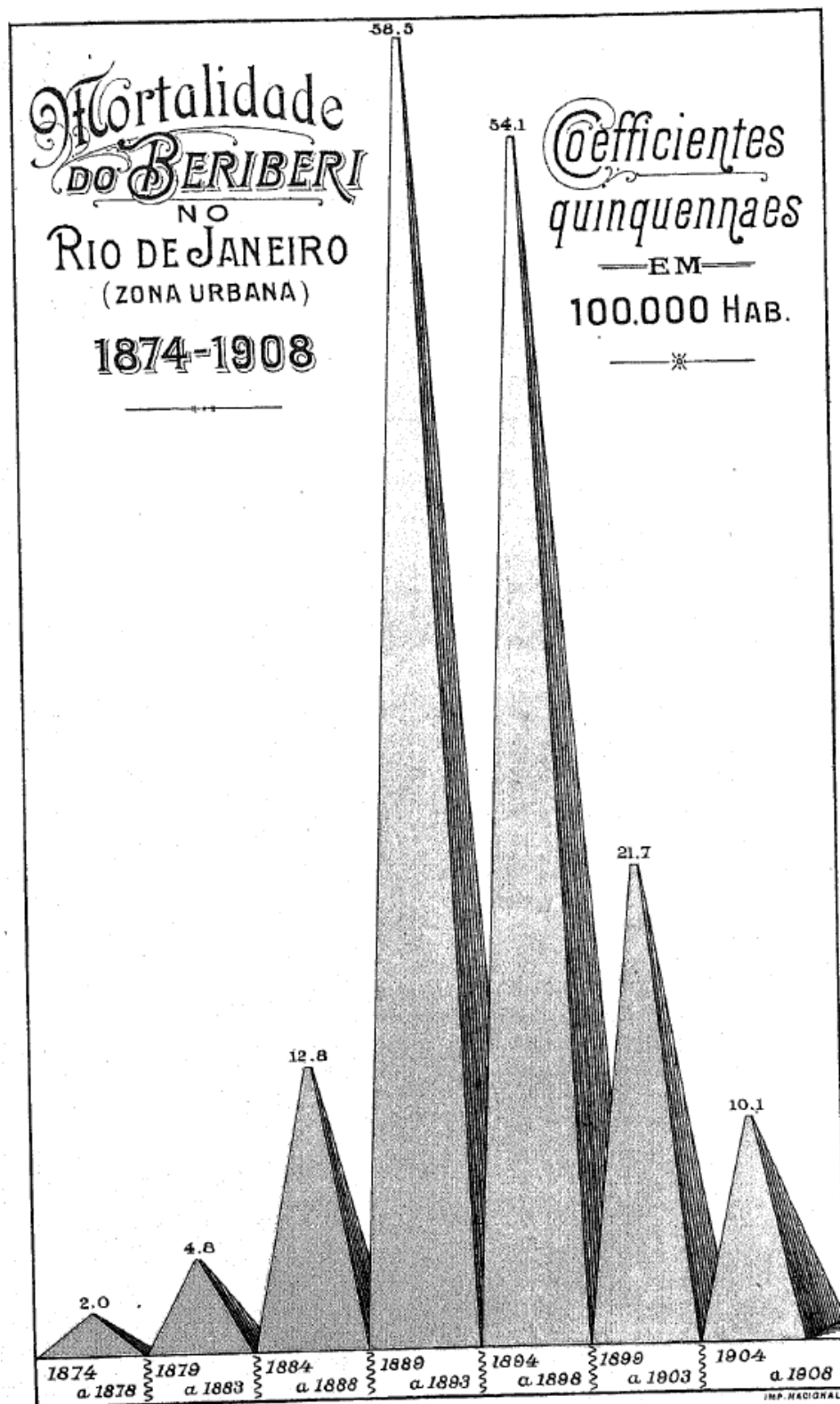
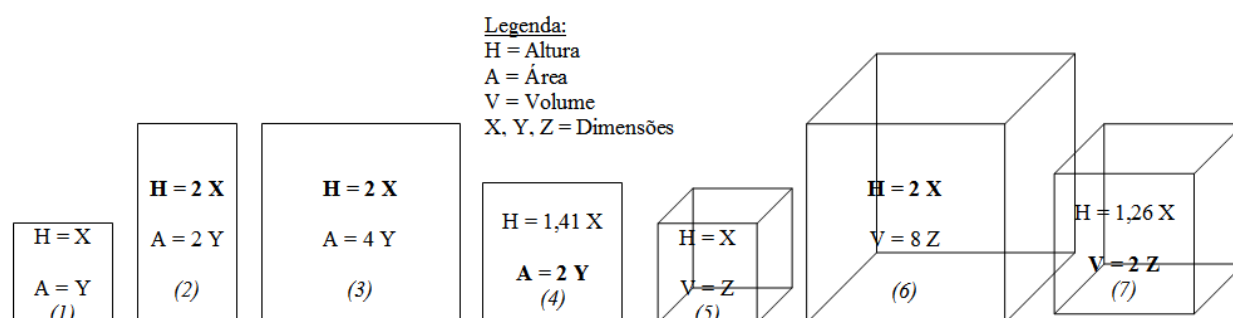



Gráfico no qual as barras são representadas por pirâmides, cujos volumes se alteram ao longo do comprimento. Duas variações de desenho para representar uma variável, que é a mortalidade causada pelo Beribéri no Rio de Janeiro entre

1874 e 1908, a cada 100.000 habitantes.[22]

O uso de **figuras substituindo as barras** distorce a nossa percepção visual da quantidade sendo representada, com o alegado pretexto de ilustrar e tornar os gráficos mais atrativos. Enquanto que geralmente é a altura da figura que está representando a medida, tanto a área quanto o volume aumentam mais rapidamente do que a altura, e sua mente percebe isso, mas isso não está de acordo com os números mostrados.

Um gráfico que mostra uma figura enorme ao lado de uma minúscula pode sugerir incorretamente uma diferença enorme, baseada no volume das figuras percebida pela sua mente; mas, no entanto, é a altura que está sendo comparada.

O uso de rótulos com valores numéricos é comum neste tipo de gráfico para mitigar os efeitos da distorção das figuras, mas estas continuam mostrando somente valores relativos aproximados. Então para quê usar as figuras em primeiro lugar? Não seria melhor mostrar os números numa tabela, uma vez que as barras perderam o seu valor informativo?

Representação de uma variável, que aumenta em duas vezes. Em cada figura, a variação desta variável é representada pela variação de uma ou mais das dimensões das formas, a partir dos desenhos bases em (1) e (5).

A variação de cada medida deve ser representada pela variação de uma única dimensão no desenho, como no caso das barras de um gráfico de barras, no qual a medida é representada pela variação na altura, e a altura e a área da barra aumentam na mesma proporção (2). Quando isto não ocorre, há distorção na representação gráfica.

Um incremento de duas vezes na altura de uma figura em duas dimensões, sem distorcê-la, significa um incremento de duas vezes na largura, e com isso, a área que é uma função quadrada das dimensões, cresce quatro vezes, dando uma impressão visual de um aumento de quatro vezes, e não de duas vezes como os números do gráfico sugerem (3).

Mas se você utilizar uma figura tridimensional, a ilusão será maior, pois ao aumentar a sua altura em duas vezes, a largura e profundidade também aumentarão nessa proporção, e o volume variando ao cubo dessas dimensões irá sugerir visualmente um aumento de oito vezes, ao invés de quatro

como no caso anterior (6).

No entanto, a variação de uma medida numa figura não necessariamente precisa ser representada pela altura, e pode ser representada também pela variação da área ou do volume. Nestes casos, a altura da figura será bem inferior a que seria se esta fosse a dimensão usada para representar a variável (4 e 7). Nestes casos o produtor deve deixar bem claro para o consumidor que a medida está sendo representada pela variação de área ou volume, para evitar que o leitor use o seu hábito de comparar a altura ao ver figuras colocadas lado a lado, como se fossem barras de um gráfico.

No gráfico a seguir é usada a figura de coluna para substituir as barras de um gráfico de barras. A variável (mortalidade por febre amarela) é representada pela altura das colunas, mas enquanto elas variam em altura, variam também em largura, e por consequência em área, passando a impressão visual de uma variação muito maior que a indicada pelos números. A coluna mais alta, referente aos anos de 1872-1876, indica 7.754 mortes, enquanto que a coluna seguinte, dos anos 1877-1881, indica 2.995 mortes, uma diminuição de cerca de 62%, corretamente representada pela altura das colunas. Mas o olho do leitor não percebe somente a altura, e sim a figura inteira, cuja variação de área sugere uma diminuição de 85%.

Gráfico mostrando a mortalidade de febre amarela no Rio de Janeiro entre 1867 a 1916. O uso da figura de coluna implica a variação da altura e largura, enquanto que a medida está sendo representada somente pela altura.[23]

O gráfico a seguir sofre do mesmo problema, com a agravante do uso de uma perspectiva que distorce mais ainda o tamanho das figuras.

Na parte superior da figura, um gráfico compara o número de brasileiros deportados da Espanha entre 2000-2003 e 2004-2007, usando uma figura de mala como se fosse uma barra. Os números mostram um aumento de 43,4% nas deportações, como indicado pela seta que sai da mala menor à esquerda e se direciona para a mala maior à direita, e como representado quase que corretamente pela altura das malas, que indicam um aumento de 50% (isso sem considerar a altura das alças, outro enfeite que confunde mais ainda o leitor).

Detalhe da figura de um passaporte aberto, cuja folha é representada em perspectiva. Sobre esta folha há o desenho de malas sendo usadas como barras, mas cuja variação da área não corresponde à variação da medida.[24]

Enquanto o que está sendo comparado é a altura das malas, o olho e a mente do leitor percebe o desenho inteiro e instintivamente compara a área das figuras, cuja diferença indica um aumento de 150%, três vezes mais que o real.

O mesmo ocorre com as malas na parte de baixo, que indicam um decréscimo de 53,9% na chegada de imigrantes ilegais brasileiros na Espanha, como é corretamente indicado pela altura das malas. Porém, as áreas sugerem um decréscimo de 76%.

Mas o uso das figuras de malas como barras não era o suficiente para envolver o leitor no clima de terminal de aeroporto. O produtor desenhou os gráficos sobre a figura de um passaporte aberto, cuja folha frontal não está paralela ao plano da imagem, mas ascende da esquerda para a direita. Esta inclinação produz uma perspectiva nas figuras, que contribui para que as malas da esquerda pareçam menores, ou pelo menos visualmente menos importantes, que as malas da direita.

Há um bom indício de que o produtor do gráfico quer valorizar a notícia, exagerando a impressão sobre o número de deportados pelo Governo da Espanha, e minimizando o número de chegada de imigrantes ilegais brasileiros, induzindo o leitor à conclusão de que o aumento de rigor das ações do departamento de imigração do governo espanhol é injustificado.

Esta conclusão pode até fazer sentido e ser o resultado final de uma análise mais detalhada do caso, porém não é tarefa do desenho do gráfico induzir o leitor a essa ou aquela conclusão. Cabe ao gráfico, e ao seu produtor, comunicar a informação da maneira mais neutra possível e evitar qualquer estratégia de desenho que distorça a representação visual.

Como discutido antes, os gráficos devem mostrar a variação nos dados, e não a variação do recipiente, ou se preferir, do desenho. Mas o uso de áreas e volumes vão contra este princípio. Usar áreas para mostrar dados unidimensionais é apenas outro modo de confundir variação dos dados com variação de desenho. Variam-se duas ou três variáveis no desenho, para representar a variação de apenas uma medida.

Por isso Tufte[25] propõe o seguinte princípio: **O número de dimensões ilustradas carregando informação (variável) não deve exceder o número de dimensões nos dados**. Não use áreas ou volumes – ou seja, mais de uma dimensão - para demonstrar dados de uma variável.

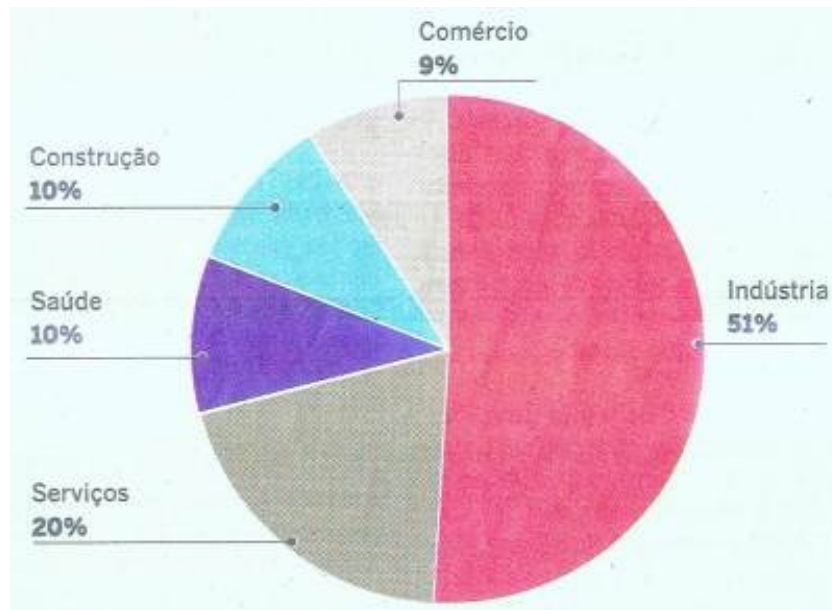
Outras armadilhas são o uso de **escalas diferentes** entre as barras tanto no eixo de medida como no eixo da base, a utilização de eixos com valores de base diferentes de zero, ou o início das barras em posições diferentes. As escalas devem ser constantes, começar na mesma posição no eixo, e as escalas devem começar no zero.

Gráficos circulares

Os gráficos circulares também são conhecidos como gráficos de pizza ou de torta. Eles são utilizados quando queremos perceber a relação entre as partes e o todo. Mas também têm suas limitações.

Primeiramente, estes gráficos só devem ser usados com medidas de proporção, nunca com valores absolutos. Além disso, eles funcionam efetivamente apenas quando as fatias são substanciais; fatias demais resultam em porções impossíveis de comparar ou até de detectar. E também, a área do círculo é uma forma inadequada para comparar quantidades simples, porque é impossível compreender pelo olhar a relação entre o diâmetro e a área de um círculo.

Hoje os gráficos de tortas são muito populares por serem muito fáceis de serem criados no computador. Ao mesmo tempo, são tão propensos a serem mal usados ou distorcerem os dados, que

Jones[26] brinca ao escrever: *"Tortas: quando estiver em dúvida, jogue uma neles!"*

Um bom exemplo de gráfico de torta, mostrando a distribuição das vagas de emprego nos setores do mercado de trabalho. Não há excesso de fatias e nem fatias muito pequenas. Este tipo de comparação é o uso ideal para este tipo de gráfico.[27]

No gráfico circular a seguir, apenas as bordas do círculo são representadas, formando o que é conhecido como gráfico de rosca. Este tipo de gráfico tem a vantagem de evitar um problema intrínseco aos gráficos de torta, que é a dificuldade visual de se distinguir a relação entre o raio do círculo, que é constante, e a área do círculo que representa o valor da variável, e varia em uma proporção do quadrado do raio, portanto não linear.

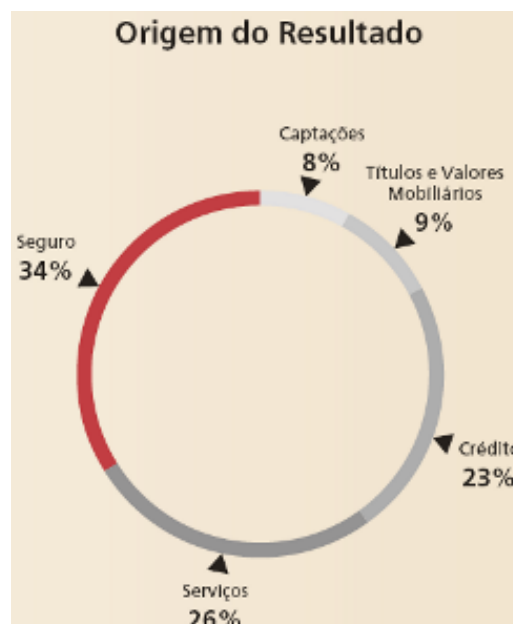



Gráfico de rosca retirado do relatório anual de um grande banco brasileiro. Um bom exemplo de gráfico: simples, sem excesso de fatias, e com rotulação direta.

Tenha em mente que os gráficos circulares são para porcentagens, razões, frações, representações decimais ou qualquer outra aplicação para mostrar a relação de proporcionalidade entre a fatia e o

todo. **As tortas nunca devem ser usadas para mostrar valores absolutos**.

Não está necessariamente errado indicar os valores absolutos juntamente com os valores de proporcionalidade, desde que estes estejam em primeiro plano. Quando as fatias das tortas são rotuladas com porcentagens ao invés de quantidades, o leitor irá focar corretamente na importância relativa das fatias ao invés de seus valores absolutos. O gráfico circular a seguir indica o valor das fatias apenas pelos seus valores absolutos, dificultando ao leitor a comparação entre as partes, o que é justamente o objetivo deste tipo de desenho.

Gráfico circular comparando a mortalidade da tuberculose com a de outras doenças transmissíveis, indicando valores

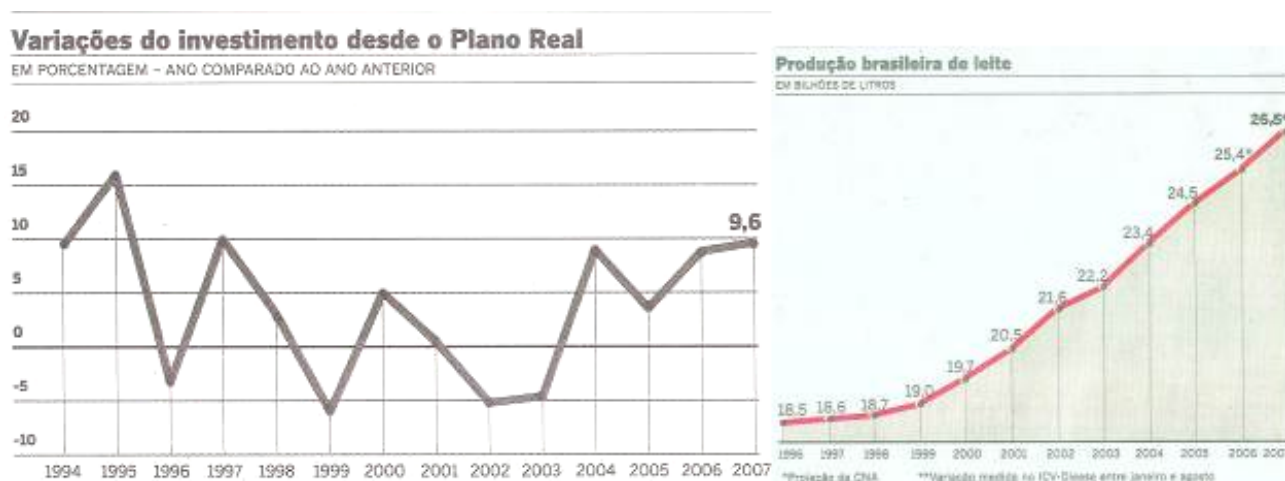
absolutos ao invés de uma medida de proporção. Valores absolutos em gráficos de torta dificultam o raciocínio.[28]

Além disso, como nos gráficos circulares os valores são sempre relativos ao total, eles podem não apresentar valores absolutos, e o valor do todo está quase sempre escondido de você; isto pode ser usado como um truque ou como uma virtude deste tipo de gráfico. Quando você se perguntar qual o valor monetário de uma fatia que indica 26% num gráfico circular sobre vendas de um produto, você não saberá.

Por essas características, é importante saber quando é melhor usar um gráfico circular ou partir para outro tipo de representação gráfica. Para comparar valores, as retas, como no gráfico de barras, são muito mais eficazes. Para comparar apenas percentuais simples, use um gráfico circular.

Além disso, para pequenos conjuntos de dados, uma tabela quase sempre é melhor do que um gráfico de torta bobo, ou pior, vários deles. As tabelas são mais simples e ocupam menos espaço que gráficos de tortas, enquanto fazem o mesmo trabalho para conjuntos de poucos dados.

Gráficos de evolução

Eles são chamados gráficos de evolução porque mostram a variação ao longo do tempo, e geralmente usam linhas, que sobem ou descem, e fazem uso da nossa percepção que para cima é mais, e para baixo é menos. Eles são ótimos para indicar tendências.

À esquerda exemplo de um bom gráfico de evolução, mostrando as variações de investimento no Brasil desde o começo do plano Real em 1994.[29] *À direita um gráfico mostrando a evolução da produção brasileira de leite; a base não zero faz parecer ao leitor desatento que a produção aumentou em 12 vezes, mas na realidade aumentou em menos da metade.*[30]

Mas assim como nos gráficos de barras, se a escala dos gráficos de evolução não for constante ou não estiver claramente indicada, o resultado será um leitor confuso ou enganado.

Os gráficos de evolução facilmente aceitam outros tipos de escala que não a linear, como as escalas logarítmicas, que são muito úteis para exibir diferenças de valores enormes com alguma precisão, mas acabam totalmente com a imaginação; é muito difícil abstrair tendências corretas ou relações

entre valores em tais escalas, a não ser que você seja muito treinado.

Tabelas

Tabelas é o tipo de representação ideal quando precisamos saber valores exatos ou representar grande quantidade de números. Nenhuma representação gráfica com escalas poderia fazer isso, a não ser que indicasse o valor numérico de cada medida ou ao longo de cada linha, o que é pouco prático.

Tabelas são muito usadas em aplicações que precisam de números exatos; os contadores e analistas precisam de tabelas em demonstrativos financeiros das empresas, economistas e investidores precisam de números exatos dos índices financeiros, e os engenheiros precisam realizar cálculos com precisão baseados em números igualmente precisos.

Além disso, para exibir grandes quantidades de números, as tabelas conseguem ocupar muito menos espaço do que se os mesmos dados fossem representados por outro tipo de gráfico.

Opções de Compra – continuação

Opções de Venda

Termo – continuação

Banco de Títulos CBLC - BTC

Detalhe de uma tabela de índices financeiros. As tabelas com os índices do mercado financeiro são exemplos comuns e efetivos de grande quantidade de informações mostradas em tabelas. No entanto o uso extensivo de abreviações e siglas faz com que somente quem conhece o mercado financeiro compreenda os seus significados.[31]

As tabelas também são indicadas para pequenas quantidades de números. É preferível usar uma tabela a ocupar espaço com um gráfico simplório. Conjuntos de dados numéricos simples devem estar em tabelas ou integrados ao texto; gráficos (que não sejam tabelas) devem ser usados para dar sentido a conjuntos de dados grandes e complexos que não podem ser representados eficazmente de outra forma.

Alguém poderia afirmar que tabelas não são gráficos, pois não fazem uso de formas geométricas, eixos ou escalas; mas a partir do momento em que elas dispõem os números de uma maneira ordenada no espaço, pode-se dizer que são uma maneira visual de ordenar a informação.

Problemas comuns em tabelas estão ligados ao seu desenho: linhas em excesso e com peso visual muito grande parecem enredar os números; falta de alinhamento entre os números; uso de resoluções diferentes para os números de um mesmo parâmetro; falta das unidades de medida; uso de abreviações em excesso. São alguns dos defeitos mais comuns.

Cuidados ao produzir gráficos

Quando você tiver que produzir gráficos, seja bom com os seus consumidores e tome alguns cuidados para não cair nas muitas armadilhas que obscurecem e confundem as informações.

O mais importante é não perder de vista o objetivo dos gráficos. É como Tufte[32] escreve: *"Representações visuais de informações devem* ***servir ao propósito analítico em questão****; se uma questão importante é uma possível relação de causa e efeito, então os gráficos devem organizar os dados de uma forma a iluminar tal ligação. Não é uma ideia complicada, mas é profunda".*

Para servir ao seu propósito analítico, a lógica que orienta a produção do gráfico tem que ter afinidade com a lógica da análise que o produtor quer induzir no leitor.

Em seguida é necessário pensar nos aspectos estruturais e de desenho dos gráficos. Primeiro, lembre-se dos **elementos básicos** de todo gráfico e verifique-os: título, legenda, eixos (nome, escala, marcações), dados e a área de fundo.

O título deve indicar claramente, e logo em primeira mão, as variáveis que estão sendo representadas. Ao invés de colocar o título "Evolução do preço do café", prefira "Preço do café (R$) x Ano de 2009 (semanas)".

Use a **rotulação direta** (aquela que rotula no próprio gráfico os valores ou a variável daquela indicação) tanto quanto possível, e use as legendas somente quando os dados forem muito complexos para a rotulação direta.

E por falar em **legenda e rotulação**, nunca se esqueça de colocá-los, assim como as unidades de medida e o que está sendo medido. Um número sem etiqueta é um número sem sentido. Geralmente uma etiqueta tem duas partes: uma unidade de medida e uma descrição do que está sendo medido. Escreva "Litros de água", "Dólares de lucro" ou "Toneladas de soja"; deve haver sempre uma unidade de medida e a descrição do que está sendo medido. Como uma regra, sempre que houver um número, deve haver um rótulo por perto o descrevendo.

Por mais óbvio que isto pareça, os rótulos são facilmente esquecidos pelo produtor que está tão familiarizado com o gráfico que não consegue imaginar que as outras pessoas não saibam o significado das cores ou siglas.

No entanto, essas legendas e rotulações não precisam ser exageradas ou ter um peso visual muito grande. O tamanho das letras em gráficos pode ser pequeno, uma vez que as frases e sentenças não são longas – e, portanto, a fonte pequena não irá fatigar o leitor da mesma forma que faria em textos longos.

Prefira sempre que possível **colocar no gráfico os dados, e não as estatísticas** desses dados. Assim evita-se a necessidade de explicar as suas suposições para chegar a tais resultados estatísticos e aumenta a eficiência de mostrar os dados. Também permite ao consumidor fazer as suas próprias conclusões, determinar se a análise do autor é apropriada e realizar as suas próprias análises, uma vez que têm os dados brutos em mãos.

Deixe bem clara a **relação de causa e efeito entre duas variáveis**. Este é outro item que pode ser negligenciado pelo produtor muito familiarizado com os próprios gráficos. Na maioria dos casos um texto acompanhando o gráfico, com explicações, análises e conclusões, é tão importante quanto o gráfico em si. E se possível, devem estar na mesma página. O gráfico na verdade não é nada mais que uma parte do texto que usa uma forma de comunicação visual, e, portanto, deve estar integrado no texto como qualquer outro parágrafo.

Use a **menor quantidade possível de marcações** na área do gráfico. Cada traço desenhado deve ter alguma finalidade informativa; se não tem, apague. Evite o já comentado "lixo de gráfico", e deixe desenhado somente o essencial. Isso significa apagar enfeites, fundos de gráfico visualmente muito pesados, linhas e números de marcação de escala em excesso, legendas demais e tudo o mais que polui e confunde.

Tenha em mente que a **seleção dos eixos para cada variável**, normalmente os eixos horizontal e vertical, irão influenciar como a informação é interpretada. Assim como a **seleção da escala**, que deve dar preferência para as escalas lineares, que são mais fáceis de compreender e comparar. Estes dois assuntos serão explicados com mais detalhe a seguir.

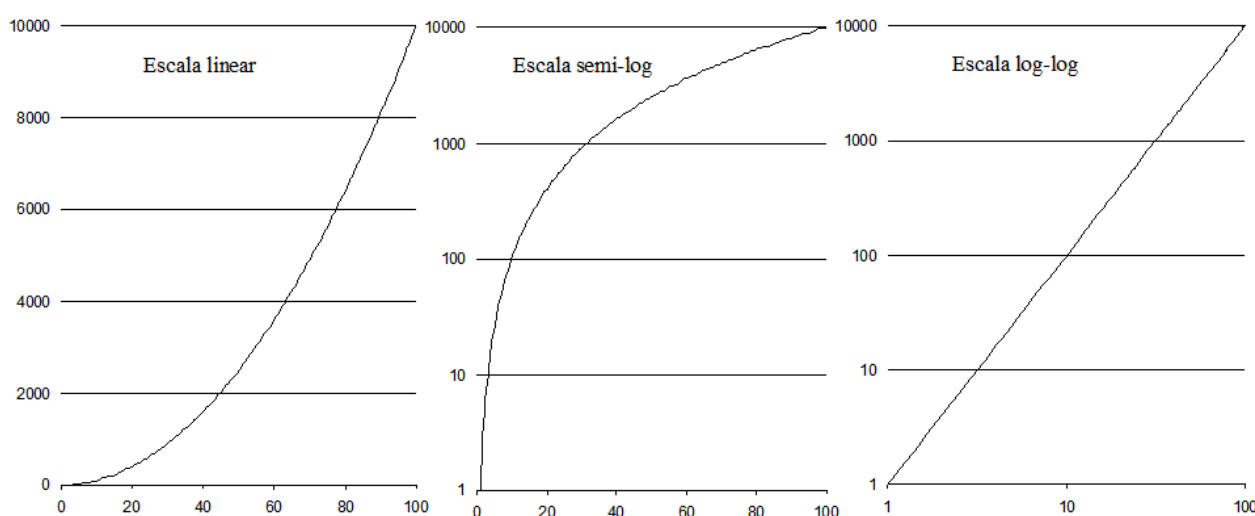
Escalas logarítmicas são traiçoeiras

Sempre que possível use escalas lineares. As escalas logarítmicas não são desonestas em si, mas por serem pouco intuitivas, podem ser facilmente manipuladas para enganar o público. O propósito honesto das escalas logarítmicas e que deve nos guiar no uso delas é possibilitar que dados de uma faixa muito larga caibam em um gráfico relativamente compacto.

No entanto, tenha em mente que este tipo de escala não é facilmente entendido pelo público em geral. As escalas logarítmicas são usadas principalmente por cientistas e engenheiros, que tipicamente lidam com dados de faixa larga, como as distâncias astronômicas, parâmetros de elementos elétricos ou estruturais, fenômenos físicos, entre outros.

Usar escalas logarítmicas é um truque para fazer curvas exponenciais (geralmente associadas com situações de crise iminentes) se tornarem lineares ou tendendo a uma estabilização. Se você tem uma curva que sobe ou desce numa proporção chocante, e informar isso vai contra os seus interesses, basta colocar o eixo desta variável numa escala logarítmica apropriada para torná-la uma reta com uma suave subida ou descida, ou mostrar uma tendência de estabilização, o que vai parecer bem menos ameaçador. Você pode transformar uma notícia catastrófica numa notícia de pouco apelo.

A escala logarítmica pode ser aplicada em apenas um eixo (cujo gráfico fica conhecido como *semi-log*), ou nos dois eixos (cujo gráfico fica conhecido como *log-log*), este último sendo muito mais difícil de ser interpretado pelo público em geral.

A sequência de gráficos a seguir ilustra o mesmo conjunto de dados desenhado com escalas diferentes. A escala do primeiro gráfico é linear e mostra que os dados apresentam uma tendência exponencial, subindo rapidamente. Caso seja de interesse do produtor, ele pode aplicar uma escala logarítmica no eixo vertical, e o mesmo conjunto de dados passa a sugerir uma tendência de estabilização. Se, no entanto, o produtor aplicar uma escala logarítmica no eixo horizontal também, os dados mostram uma reta, que parece bem mais familiar e menos chamativa.

O mesmo conjunto de dados desenhados em gráficos com escalas diferentes. À esquerda, uma escala linear mostrando uma curva exponencial; ao centro, um escala logarítmica no eixo vertical que mostra os dados com uma tendência de estabilização; à direita, escalas logarítmicas em ambos os eixos mostram os dados como uma reta.

Para fazer um uso honesto das escalas logarítmicas, certifique-se de indicar com clareza quais eixos

estão usando-as.

Considere os preconceitos culturais sobre as orientações

O produtor de gráficos deve ter em mente que as pessoas têm preconceitos e tendências culturais que as fazem intuitivamente dar certos significados às direções adotadas nos gráficos. A maioria das pessoas nem tem consciência de ter essas polarizações culturais, e mesmo as que sabem disso, não conseguem se desvencilhar dessas tendências tão arraigadas.

O bom produtor de gráficos deve estar ciente dessas polarizações e deve saber usá-las para reforçar a sua mensagem; ir contra elas apenas prejudicará a comunicação.

Para Jones[33], essas polarizações culturais surgem principalmente da direção de leitura de cada cultura. Para os ocidentais, ele as descreve assim:

"*"As pessoas nas culturas ocidentais leem da esquerda para a direita. Para estes leitores o movimento para a **direita** – a direção em que o olho varre a página – é associada com a passagem do tempo, e, portanto, com o movimento positivo, ou mesmo com a ideia de **progresso**"*."

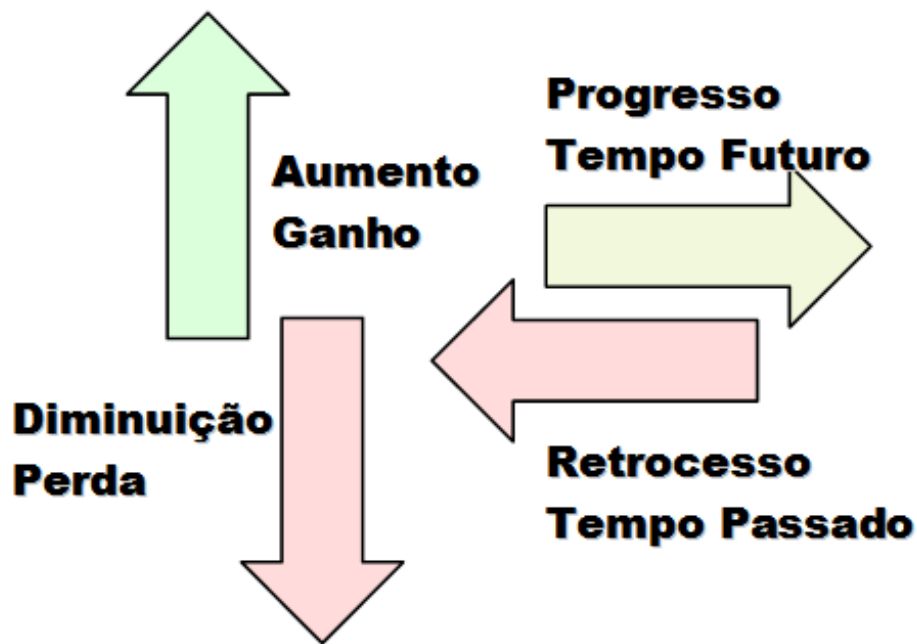
"*"Inversamente, o movimento da direita para a **esquerda** é considerado **retrogrado** e negativo – mau! Dessas mesmas noções vem a raiz da palavra **sinistro**, que pode significar tanto esquerdo ou de mau agouro."*"

Os gráficos que mostram a passagem do tempo, como os de evolução, usam essa noção que o tempo se move da esquerda para a direita.

Estas polarizações culturais também se revelam na etimologia das palavras, o que dá indícios sobre a sua antiguidade e enraizamento cultural. Sinistro é derivado de *sinister* em latim, e pode significar tanto "esquerda" ou "à esquerda" quanto "mau", "azarado", "pernicioso" e também é usado no mercado de seguros para referir-se a acidentes ou prejuízos materiais. O seu antônimo, destro, deriva de *dexter* em latim, e pode significar tanto "direita" ou "à direita" quanto "correto", "afortunado", "habilidoso", "próprio" e outros adjetivos positivos. Como várias línguas são herdeiras do latim, esta significação foi passada adiante e hoje é comum a muitos idiomas, sendo os mais comumente falados o inglês, francês, espanhol e italiano, além do português.

Sobre os movimentos verticais, Jones[34] escreve:

"*Em muitas culturas, o movimento para cima é associado com aumento ou ganho, e para baixo com diminuição ou perda.*". Portanto: "Gráficos que mostram flutuações em quantidades se baseiam na noção de que ***para cima*** significa **ganho** e ***para baixo*** significa **perda**."

Esta noção também vem da direção em que as pessoas costumam ler os textos, que é de cima para baixo.

As quatro direções e suas conotações de acordo com a cultura ocidental. Essas conotações mudam com as culturas e podem influenciar na maneira de interpretar um gráfico.

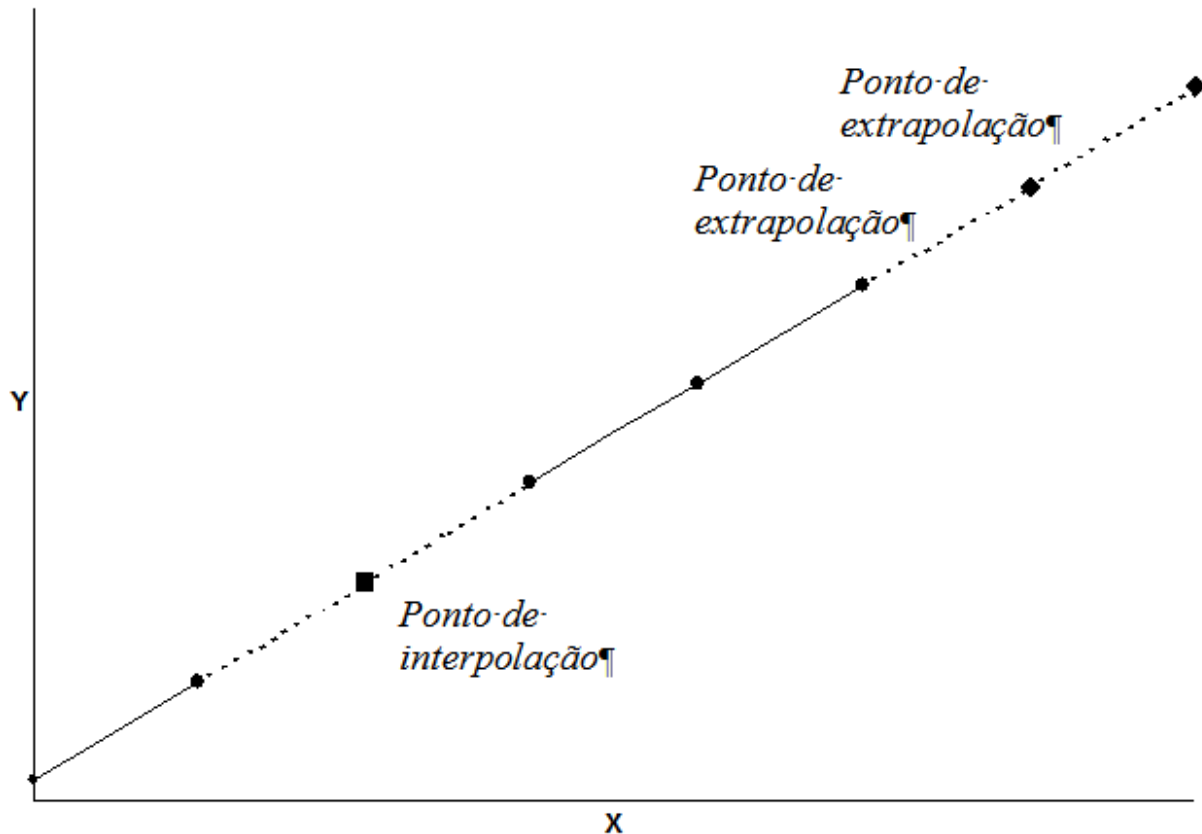
Assim, nas culturas ocidentais, as pessoas intuitivamente entendem que *para cima* e *para baixo* significa *quantidades* e para a *direita* e para a *esquerda* significa *tempo*.

No entanto, essas predisposições para dar conotação às direções variam de acordo com a cultura e a direção de leitura de cada língua. Para quem lê hebreu ou árabe, cuja direção de leitura é da direita para a esquerda, um gráfico que mostra a passagem do tempo para a esquerda (o contrário dos ocidentais) pode parecer mais familiar. Já para os chineses que leem de baixo para cima, as flutuações verticais de quantidades podem não parecer tão intuitivos assim.

Mas se você não estiver atento, mentirosos espertos poderão usar essas predisposições para conotar sentido às orientações, e assim comunicar um sentimento oposto ao que os dados transmitiriam. Você pode fazer um gráfico de lucros despencando não parecer tão mau assim se você inverter o sentido de fluxo do tempo. Um leitor desatento pode até acreditar que os lucros estão na verdade aumentando.

Adivinhando os dados que faltam

Gráficos são meios eficientes de mostrar padrões e tendências em conjuntos de dados. A representação visual de linhas ou outras formas geométricas que parecem seguir determinada direção ou forma é bastante persuasiva para que o produtor se sinta seguro, ou mesmo considere

óbvio, o preenchimento de dados faltantes ou a projeção de dados futuros. Mas qualquer tentativa de expandir um conjunto de dados por meio de adivinhação de valores internos (processo chamado de **interpolação**) ou por identificar tendências e projetar valores futuros (processo chamado de **extrapolação**) é perigosa e sujeita a vários truques ou erros não intencionais.

Gráficos lineares podem ser facilmente completados com pontos de ***interpolação*** *(valores internos) e* ***extrapolação*** *(valores futuros). Entretanto, poucas situações apresentam tal simplicidade, exigindo do produtor métodos mais complexos para determinar estes pontos.*

O problema começa quando o desenhista de um gráfico faz interpolações ou extrapolações baseadas em suposições erradas ou tendenciosas, por má fé ou uma inocente falta de habilidade, mas quase sempre convenientes ao resultado ao qual o produtor quer chegar. Se as suposições para um processo de interpolação ou extrapolação são escondidas ou omitidas do leitor, desconfie.

Na prática, a extrapolação e a interpolação são ferramentas úteis somente se os fenômenos observados podem ser descritos com precisão por fórmulas, que são usadas para definir os pontos em questão.

Como exemplos de situações que possuem fórmulas bem definidas e que podem fazer um uso honesto da interpolação ou extrapolação, estão: aplicações financeiras com rendimentos conhecidos, alguns comportamentos da economia, leis da natureza, alguns fenômenos biológicos, entre outros.

Mudanças na escala horizontal ou na largura afetam a percepção

A escolha da largura de um gráfico afeta a percepção da evidência dos dados. Gráficos estreitos e altos ajudam a enfatizar as variações. Gráficos largos ajudam a amenizar as variações. O produtor pode escolher a razão entre altura e largura que melhor conduzir ao significado que ele quer passar.

Se o produtor tem um gráfico que mostra muitas flutuações, ou tendências indesejadas de subida ou queda, basta alongar o gráfico para amenizar tais evidências visuais. Caso queira enfatizar tais variações, basta tornar o gráfico alto e estreito.

Um efeito parecido pode ser conseguido com outro truque mais sutil, que é mudar o valor da escala no eixo horizontal, sem nenhum razão aparente. O leitor incauto não perceberá as escalas que variam de valor, e será conduzido somente pela evidência visual do gráfico.

O gráfico a seguir usa este truque de uma maneira bem leve. O tempo entre a primeira barra e a terceira barra equivale a dez anos; logo, o leitor supõe que a barra do meio, colocada igualmente espaçada entre as duas outras barras, representa um valor do meio, ou seja, de cinco anos; no entanto, a barra do meio representa seis anos. Neste caso, para passar uma evidência visual correta, esta barra deveria ser colocada ligeiramente para a direita, mais próximo à barra dos dez anos. O observador que não se atenta a essa mudança, tem a impressão de uma tendência de queda menor do que realmente é.

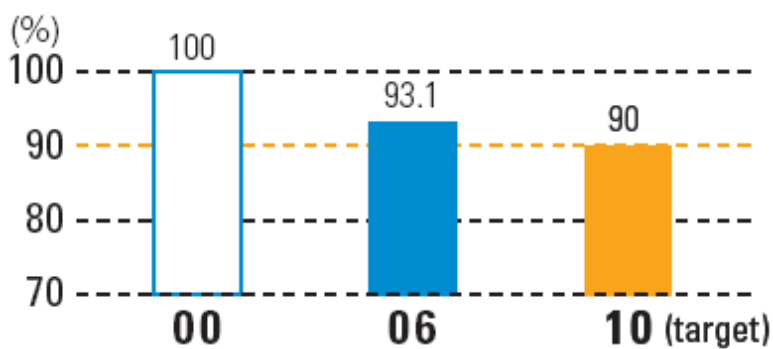



Gráfico retirado do relatório anual de uma grande companhia automobilística japonesa. As distâncias na escala horizontal não coincidem com os valores de tempo, e passam uma impressão de variação distorcida ao leitor.

O gráfico a seguir é um ótimo exemplo de como não fazer um gráfico, ou como cometer vários erros de representação visual em alguns centímetros quadrados.

Um dos seus problemas é bem comum, que é a escala que não começa no zero, o que dá uma impressão exagerada na taxa de mudança dos valores. Enquanto parece que o valor de mercado das empresas americanas despencou de um pico a zero (isso mesmo, o gráfico dá a impressão que as empresas não valem mais nada, uma vez que o final da linha de evolução encosta no eixo horizontal), na verdade elas desvalorizaram cerca de 17% (de 17,054 para 14,082 trilhões de

dólares).

Mas o problema mais grave está na escala horizontal, que divide o tempo analisado em três períodos: março a outubro, outubro a dezembro, e dezembro a março. Entretanto, as divisões não apresentam tamanhos iguais, e nem denotam o mesmo período de tempo.

A primeira divisão representa 6 meses (abril, maio, junho, julho, agosto e setembro), e tem 80% do tamanho da segunda divisão. A segunda divisão representa apenas 1 mês (novembro) e é maior que a primeira, que representa 6 meses, e ligeiramente maior que a terceira, que representa 2 meses (janeiro e fevereiro).

Caso este gráfico estivesse representado com um escala de tempo coerente e espacialmente proporcional ao tempo representado em cada período, poder-se-ia notar uma tendência de queda muito mais acentuada que a mostrada por este, o que corroboraria o ponto de vista da reportagem, que tenta chamar a atenção para a crise da bolsa de valores americana.

Este é provavelmente um caso em que o desenho de gráfico ruim atuou contra os interesses do seu produtor, amenizando, ou invés de exagerar, as evidências da notícia. Nestes casos pode-se afirmar, com certa dose de confiança, que o produtor sofreu de desinformação, falta de atenção ou ignorância.

Gráfico com problemas nas escalas vertical e horizontal. Na medida do tempo, não há coerência ao longo do eixo sobre a proporcionalidade entre a medida usada no desenho e o tempo representado.[35]

Mas o uso de mudanças nas escalas não é intrinsecamente incorreto e nem sempre é desonesto. Desde que a mudança na escala seja claramente indicada ao leitor, e este esteja ciente de que a representação visual pode estar distorcida, não há problemas.

A seguir, no eixo vertical da direita (que indica porcentagens), há uma indicação de descontinuidade na escala. O símbolo de descontinuidade (≈) é uma solução muito mais honesta do que

simplesmente começar a escala num valor diferente de zero.

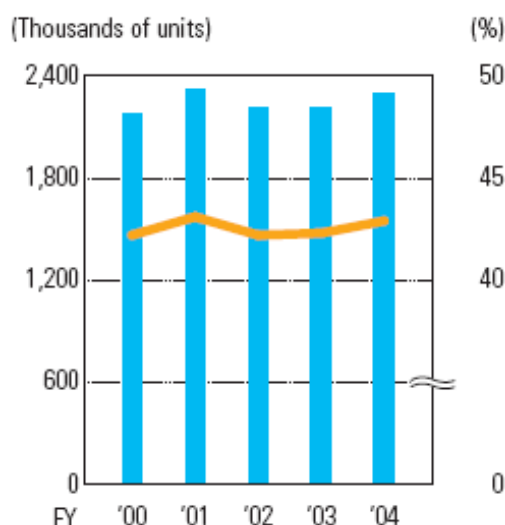



Gráfico retirado do relatório anual de uma empresa automobilística japonesa, com uma indicação de descontinuidade na escala do eixo vertical direito. Um exemplo de uso honesto de escala descontínua.

O próximo exemplo faz um uso mais extensivo das descontinuidades, e ainda assim de forma honesta. A escala horizontal apresenta o símbolo de descontinuidade para enfatizar ao leitor que há uma diferença de escala entre os anos de "90" e "02", que logo é seguido por anos consecutivos. Esta indicação é repetida na linha de evolução na área do gráfico, reforçando a indicação de mudança de escala. O mesmo acontece com a escala vertical, que apresenta duas linhas em paralelo que vão serpenteando através das barras para indicar uma descontinuidade na escala, que corretamente começou no zero. Um produtor desonesto teria usado uma escala de base não zero, ou mesmo esta descontinuidade, sem qualquer preocupação em indicar ou chamar a atenção sobre isto.

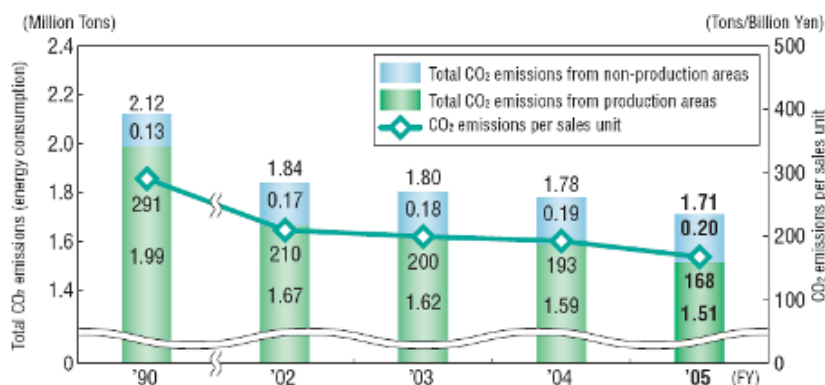



Gráfico retirado do relatório anual de uma empresa automobilística japonesa, com indicações de escalas descontínuas em ambos os eixos. Escalas descontínuas podem ser usadas honestamente, desde que indicadas ao leitor.

Usar escalas descontínuas sem evidenciar isto ao leitor é desonesto, pois fornece uma representação visual distorcida em relação ao que o leitor está esperando do desenho do gráfico.

Mas mesmo o uso de escalas descontínuas indicadas ao leitor deve ser usado com cuidado, pois mesmo sendo tecnicamente correto e honesto, continua distorcendo a imagem geral do gráfico e

dificultando a análise. O leitor é obrigado a abstrair mentalmente como seria o gráfico se a escala fosse contínua, para ter uma noção mais realista das tendências e variações dos valores. Além disso, as descontinuidades nas escalas geram algumas perguntas incômodas: O que o autor está deixando de fora? O que há na área descontínua que poderia mudar a análise do gráfico? Se a escala fosse contínua o gráfico transmitiria outra mensagem?

Gráficos empilhados podem esconder informações

Gráficos de áreas ou barras empilhadas é uma boa maneira de aglomerar vários conjuntos de dados numa mesma área de desenho. Porém eles também têm potencial para enganos e mentiras.

A primeira pergunta que vem é: as barras ou áreas estão colocadas uma sobre as outras, ou uma atrás das outras? Jones[36] explica bem este problema: *"As áreas ou barras estão empilhadas verticalmente – como tijolos e argamassa – ou estão em camadas uma em cima do outra – como desenhos em folhas transparentes separadas, e cada nova área começando na mesma linha horizontal na parte de baixo do gráfico? Se as áreas estão realmente empilhadas, somente os dados de baixo terão uma linha de base plana. Cada nova área terá o topo da área precedente como a sua linha de base, e qualquer flutuação nesta linha irá distorcer os altos e baixos da área seguinte."*

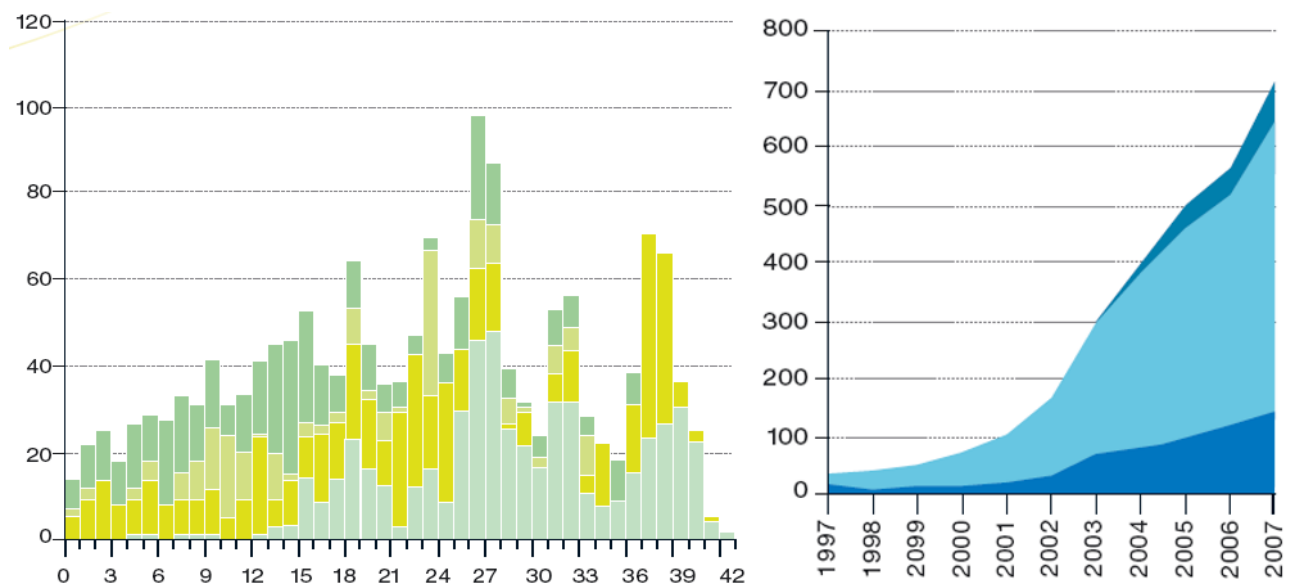



Gráfico de barras empilhadas (à esquerda), e de áreas empilhadas (à direita), retirados do relatório de um órgão governamental europeu. Note que somente o primeiro conjunto de dados, aquele cujas barras ou áreas se iniciam no eixo horizontal, é que apresentam a mesma linha de base.

Se os dados estiverem realmente empilhados como tijolos, significa que a o conjunto de dados seguinte usa como base a parte superior do conjunto de dados anterior. Portanto, é claro que quanto maior a variação do dado que é colocado na parte de baixo, maior será o efeito de flutuação do dado que está em cima. Este efeito pode ser minimizado pelo produtor bem intencionado, que coloca

perto da base o conjunto de dados com menor variação, e, portanto que causará menos flutuação aos conjuntos acima. Mas caso o produtor queira gerar confusão ou confundir o leitor, poderá colocar a variável com maior flutuação na parte de baixo, e transformar a leitura do gráfico em uma árdua tarefa.

Já os gráficos que empilham os conjuntos de dados em camadas não apresentam este problema, pois todos os conjuntos de dados começam na mesma base, o eixo horizontal. No entanto, só funcionam bem se o conjunto de dados à frente apresenta sempre valores menores do que os conjuntos de dados atrás; caso em contrário, partes dos conjuntos de dados de trás ficarão escondidos, o que pode até ser conveniente para um produtor desonesto.

O gráfico a seguir é um bom exemplo deste tipo de representação em camadas.

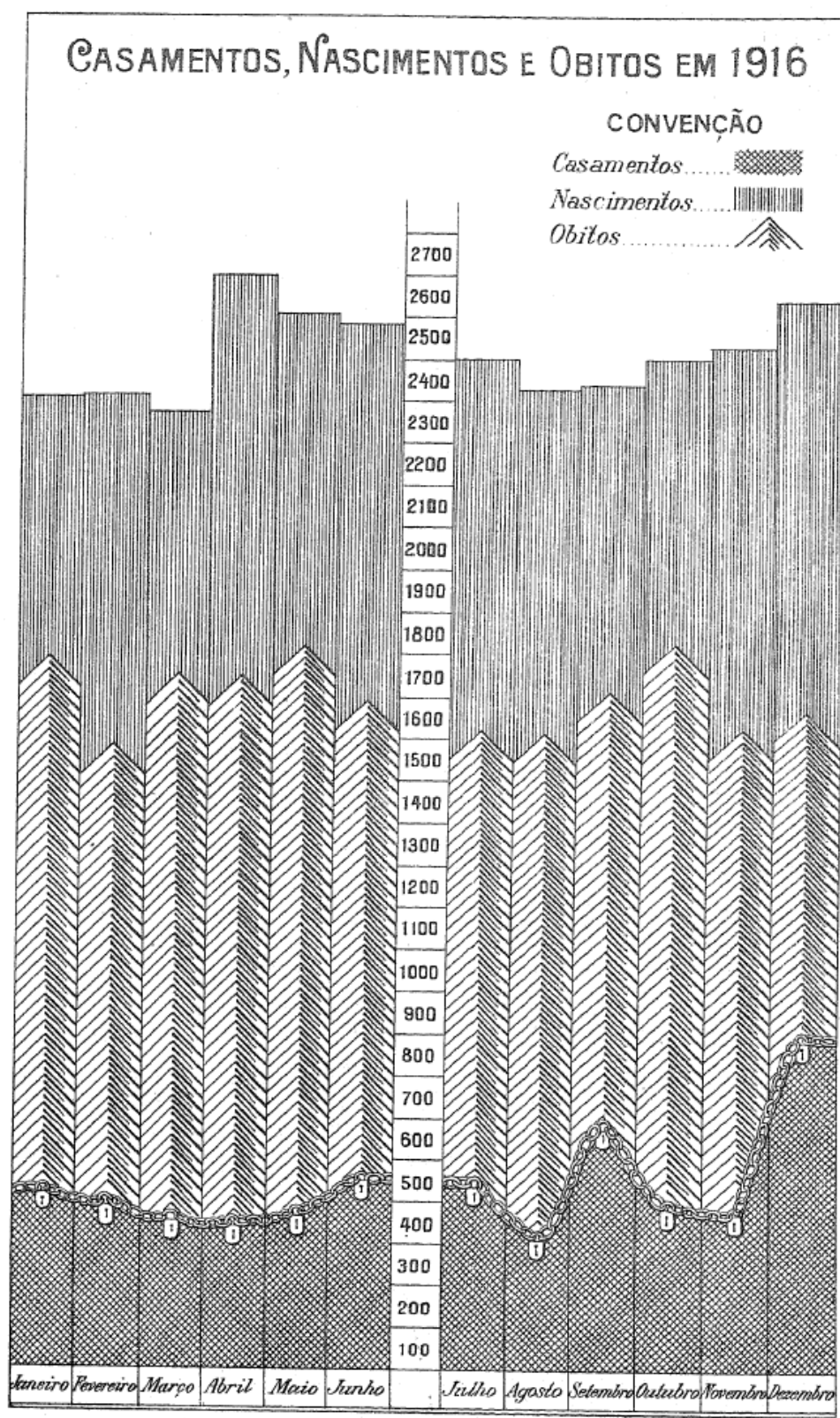



Gráfico com as estatísticas de nascimentos, óbitos e casamentos em 1916, a cada mês. Um bom exemplo de gráfico empilhado em camadas, apesar da representação visual pesada gerada pelas hachuras. Além disso, o desenho não é neutro ao usar uma corrente e cadeados para demarcar a linha dos casamentos, e poderia gerar polêmica.[37]

A camada anterior é a de nascimentos, a camada seguinte é a de óbitos, e a camada mais frontal é a de casamentos. Esta ordem foi uma escolha lógica, pois como geralmente uma população apresenta crescimento, o número de nascimentos será maior que o de óbitos. Já o número de casamentos

implica a consideração de duas pessoas, o que quase sempre irá significar um número menor que o de nascimentos e óbitos (as pessoas se casam em pares, mas nascem e morrem individualmente).

Gráficos sobre dinheiro só fazem sentido se os valores têm a mesma referência

O valor do dinheiro muda com o tempo devido à inflação. Como a inflação é quase sempre positiva, o dinheiro tende a desvalorizar com o tempo. Não faz sentido comparar o valor nominal atual e de vinte anos atrás do preço de um quilo de laranja ou de um quilo de carne. No entanto, faz sentido comparar o preço entre as laranjas e a carne hoje, e entre as laranjas e a carne há vinte anos, pois embora sejam produtos diferentes, apresentam a mesma base monetária, o que permite uma comparação de preços.

Um gráfico que apresenta os **valores nominais** dos preços no tempo apresenta além da variação da medida (o preço), uma variação no desenho (preços com valores reais diferentes). A escala vertical efetivamente muda, pois o valor do dinheiro muda com o tempo, distorcendo o gráfico. Então, o único modo de pensar claramente sobre dinheiro no tempo é fazer comparações usando unidades monetárias ajustadas pela inflação. Ajuste os valores do passado de acordo com a inflação do período, e assim você saberá que está apresentando os preços em valores reais, e não nominais.

Um exemplo que está constantemente nos noticiários é o preço do barril de petróleo. Os jornalistas estão a todo o momento anunciando, com um semblante de preocupação, a quebra de novos recordes no preço do barril de petróleo. Considerando que esse preço é cotado em dólares americanos, e há inflação nos Estados Unidos, então é de se esperar que com o passar do tempo o preço nominal seja reajustado e chegue a patamares inéditos, no entanto, sem implicar necessariamente num aumento real.

O gráfico a seguir mostra a variação do preço do barril de petróleo de Janeiro de 2000 a Janeiro de 2005. O observador perceberá uma clara tendência de aumento, e logo concluirá que deverá esperar por um aumento no preço dos combustíveis em breve.

O que pode passar despercebido é que o título do gráfico indica que estes são os valores nominais, e, portanto, sem o ajuste pela inflação. Se este gráfico tiver os valores do passado ajustados pela inflação, eles tornar-se-ão valores reais, o que geralmente significa valores absolutos maiores. Quanto mais no passado estiver o valor, maior a inflação acumulada no período, e maior será a diferença entre o valor nominal e o real.

O gráfico com os valores reais tenderá a ser mais plano, e a tendência de subida dos preços ficará menos evidente.

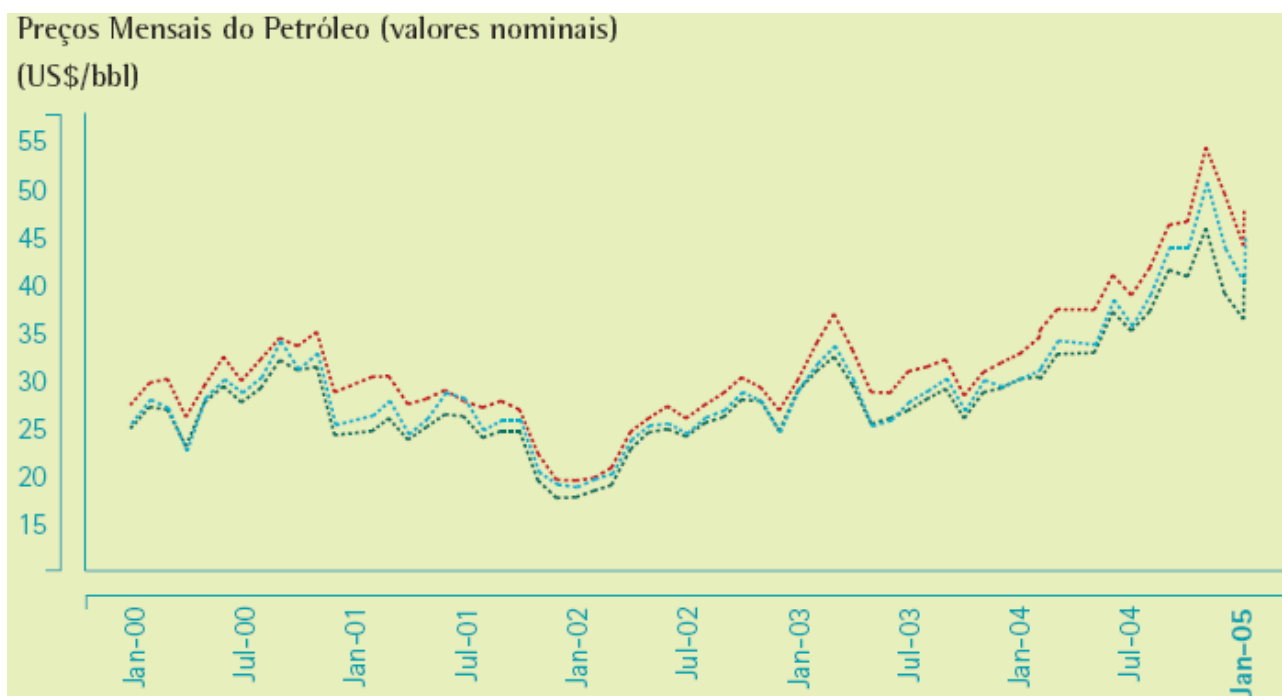



Gráfico retirado do relatório anual de uma grande estatal brasileira, mostrando a variação do preço nominal em dólares americanos por barril do petróleo entre Janeiro de 2000 e Janeiro de 2005, de acordo com três cotações. Valores de dinheiro no tempo não deflacionados causam uma variação no desenho, pois o eixo vertical muda.

Tufte[38] sugere o seguinte princípio para gráficos sobre dinheiro: ***"Em visualizações de dinheiro no tempo, unidades de medida monetárias deflacionadas e padronizadas são quase sempre melhores que unidades nominais."***

Os gráficos orçamentários dos governos também devem considerar os efeitos da inflação, assim como uma segunda fonte de distorção: a variação da população. As medidas orçamentárias dos governos, como a arrecadação de impostos, as despesas com serviços públicos ou os investimentos em determinado setor, são diretamente proporcionais à população que paga impostos e utiliza os serviços públicos.

Do mesmo modo que a comparação entre valores monetários em tempos diferentes requer a utilização de uma unidade padronizada de medida (como, por exemplo, os preços deflacionados ou reais), a comparação de receitas e despesas do governo só faz sentido quando reduzidas a uma unidade padronizada de medida que leve em consideração a população considerada. **Em gráficos governamentais, faz muito mais sentido pensar em termos de unidades monetárias reais *per capita***.

Os gráficos de orçamentos e gastos governamentais geralmente dão a impressão de estarem subindo rapidamente, o que serve de material para muitas reportagens sensacionalistas e críticas ao governo. Entretanto, é de se esperar que os valores de arrecadação e gastos dos governos constantemente cheguem a valores nominais inéditos, tanto pelo efeito da inflação, quanto pelo aumento da

população. Este efeito só não ocorrerá em localidades com deflação ou decréscimo populacional.

Assim, enquanto gráficos em valores nominais podem indicar um aumento de orçamento ou despesas, quando normalizados pela inflação e pelo aumento da população, podem indicar até uma queda.

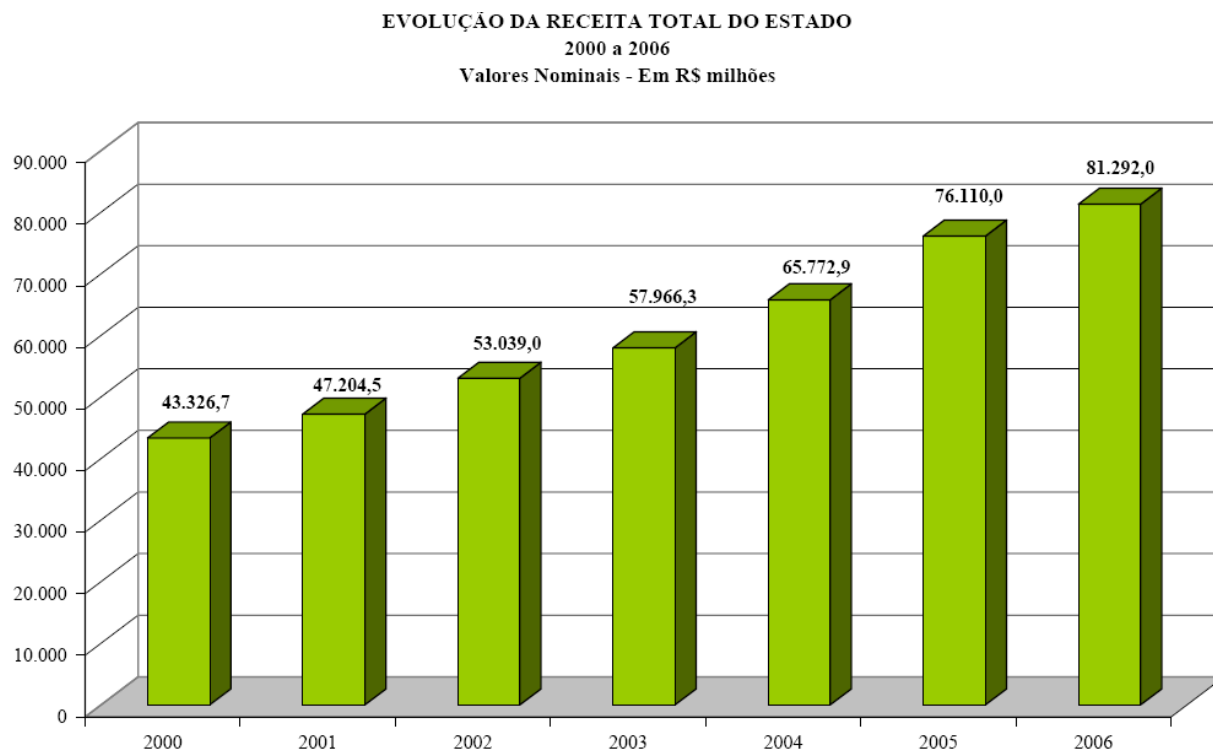



Gráfico de evolução da receita do Governo do Estado de São Paulo entre 2000 e 2006, que indica um crescimento estabilizado. Note que os valores apresentados são nominais. Se as medidas deste gráfico fossem padronizadas em unidades monetárias reais per capita, poderiam indicar um nível constante ou até mesmo um decréscimo de receita.[39]

A aplicação da aritmética faz possível levar em consideração a população e a inflação. Computar gastos em unidades monetárias reais per capita geralmente revela uma situação bem diferente e mais precisa do que os gráficos de medidas não padronizadas.

Uso de gráficos 3D

Os gráficos em três dimensões são extremamente fáceis de elaborar no computador; tão fáceis quanto qualquer outro gráfico em duas meras dimensões. Então porque continuar usando os tradicionais e sem graça gráficos em 2D, se podemos usar gráficos em 3D que parecem muito mais legais e passam uma impressão de modernidade e sofisticação? Simples, porque para a grande maioria dos gráficos, não é necessário o uso de três dimensões e o seu uso só adiciona complexidade e confunde o leitor. Use gráficos em 3D somente se a terceira dimensão adicionar alguma informação.

Como Wurman[40] descreve: *"A possibilidade de os gráficos mostrarem profundidade cria mais confusão quando esta profundidade não tem valor numérico. Isso acontece tanto em gráficos*

circulares, como de barras, e bem como em mapas geográficos. Espera-se que você olhe o plano ou a profundidade? Isso serve apenas para confundir a sua percepção da informação."

E Jones[41] é ainda mais enfático ao criticar os gráficos em 3D: *"Quando você precisar escolher um tipo de gráfico no PowerPoint, se você quiser enganar e distorcer, escolha tortas ou barras dimensionais. É virtualmente impossível fazer gráficos 3D que não dão a impressão errada."*

Os gráficos em 3D também sofrem do problema que Wurman chama de "despenhadeirite", que é essa profundidade adicionada aos gráficos, que na maioria das vezes parecem como despenhadeiros circundando-os, como no exemplo a seguir. São ainda mais graves quando aparecem em mapas 3D, pois passam a impressão de um verdadeiro despenhadeiro, que poderia ter muitos quilômetros de profundidade se medido na mesma escala do mapa.

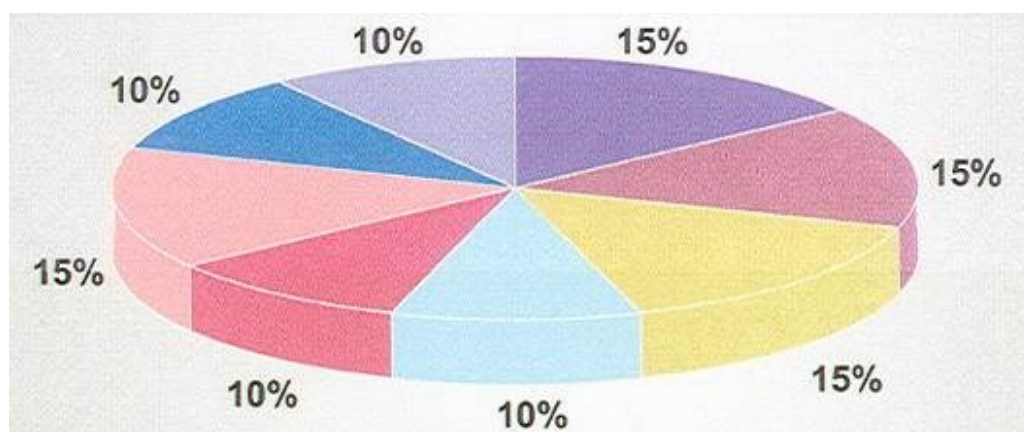



Gráfico de torta em 3D retirado do prospecto de fundos de investimento de um grande banco brasileiro, mostrando a composição de determinada carteira. Exemplo do uso desnecessário de uma perspectiva em 3D em um gráfico simples, que passa a sofrer de "despenhadeirite."

Quando você olha o gráfico, sabe que a profundidade não está representando nenhuma informação, mas a sua mente vê essa profundidade e tenta tirar algum significado disso ou o compara com as outras medidas. O uso de uma dimensão a mais que não traz nenhuma informação numérica só distrai e confunde. O uso de gráficos 3D quando não são necessários é mais um exemplo de "lixo gráfico", que mostra uma estupidez gráfica e denigre a imagem do produtor.

Mas os gráficos em 3D podem ser muito úteis quando apropriadamente usados, e não é necessário nem um computador para isso; uma prova que somente ferramentas excelentes não fazem gráficos igualmente excelentes, e a responsabilidade pela sua qualidade ainda recai sobre os ombros do produtor.

O estatístico italiano **Luigi Perozzo** ficou conhecido na segunda metade do século XIX por seus excelentes diagramas em 3D. Em 1879 ele publicou o seu diagrama mais conhecido, uma representação da pirâmide da idade da Suécia, baseada nos dados estatísticos do censo sueco de

1750 a 1875. O gráfico mostra a evolução da pirâmide de idades da população sueca entre 1750 e 1875. O eixo vertical (altura) representa a quantidade de pessoas, o eixo longitudinal (profundidade) representa as faixas etárias e o eixo lateral (largura) representa os anos, começando em 1750 e terminando em 1875.

A primeira linha no eixo de idades (mais ao fundo) representa os nascidos vivos, e ela se sobressai sobre o resto do gráfico. Podem-se notar dois fatos interessantes com essa linha: o primeiro fato é que a faixa etária seguinte apresenta uma queda expressiva em população, denotando uma alta mortalidade infantil, mas que tende a diminuir com o passar dos anos; o outro fato é a ascensão rápida do número de nascidos a partir de 1850, o que provoca um degrau na pirâmide dos anos seguintes. Ambos os fatos poderiam ser atribuídos à melhoria da qualidade de vida e aos avanços da ciência e da medicina.

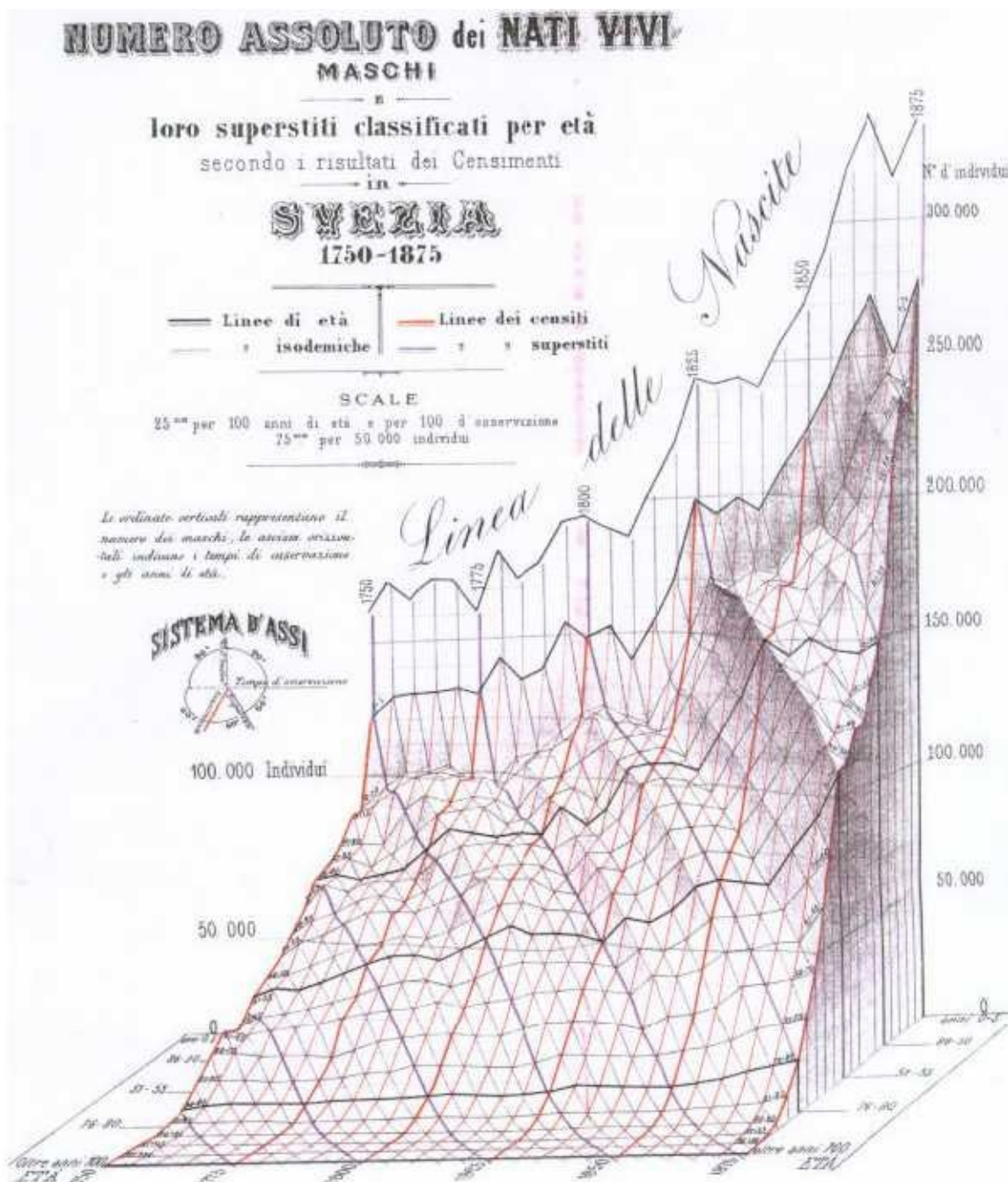



*Gráfico em 3 dimensões de **Luigi Perozzo**, 1880, mostrando a evolução da pirâmide de idades da população da Suécia entre 1750 e 1875. Um exemplo precoce de um bom gráfico em 3 dimensões.*[42]

Relatórios anuais e gráficos ruins

Os relatórios anuais de empresas são divulgados para exibir o seu desempenho econômico e financeiro para investidores e analistas de mercado. Recentemente também estão divulgando os relatórios sociais, nos quais são descritas as ações da empresa para a melhoria da comunidade e preservação do meio ambiente. A maioria apresenta um estilo sofisticado e colorido, com muitas fotos de setores das fábricas, produtos, funcionários sorridentes e grupos de diretores em ternos caros e sorrisos confiantes; mas poucos se preocupam em ser ferramentas eficazes de comunicação com sócios, acionistas e analistas de mercado.

Esses documentos geralmente seguem o mesmo estilo de apresentação de números e gráficos ao longo de todo o relatório, embora a maioria das empresas não mantenha o mesmo estilo entre os relatórios de um ano e outro, forçando os seus leitores a se adaptarem a um novo modo de apresentar as informações a cada ano. São poucas as empresas que definem um estilo visual e o mantêm, apresentando uma coerência; a maioria parece que contrata um projetista de relatório diferente a cada ano, que decide projetar o relatório da maneira que acredita ser a mais adequada.

Estes relatórios corporativos, assim como relatórios governamentais que também serão usados como exemplos aqui, estão cheios de números e gráficos, que não por acaso são de grande interesse para o próprio produtor da informação. Certamente estes relatórios se encaixam na categoria em que o produtor da informação tem grande interesse nos resultados. Só isso já é o suficiente para atribuirmos uma boa dose de desconfiança aos seus números e gráficos, possíveis portadores de distorções e apresentações tendenciosas. Especialmente se a empresa em questão não tem bons números para mostrar, e está ávida para passar uma boa imagem, ou o governo quer justificar um aumento de impostos ou exagerar os progressos realizados.

É importante ressaltar que a precisão, a clareza e a coerência da representação visual dos dados é que estão sendo avaliados aqui, e não o estilo ou a aparência dos gráficos. Certamente muitos gráficos que têm uma representação visual correta poderiam tornar-se mais atrativos ou com uma leitura mais fácil e agradável caso adotassem outras estratégias de desenho, como as recomendadas por Tufte, mas isto não tira o seu mérito por informar corretamente e honestamente o leitor.

Caso 1

Gráficos que não começam no zero muitas vezes são propositalmente usados para enganar, pois exageram as variações. O gráfico abaixo do relatório anual de uma grande empresa brasileira ilustra a parcela do mercado a qual a empresa ocupa. Um olhar de relance no gráfico sugere que a sua parcela de participação no mercado cresceu cerca 2,5 vezes ou 250% de 2004 a 2006, mas prestando atenção na escala percebe-se que aumentou de fato cerca de 4% em relação a 2004. Este é um tipo de truque usado para enganar e confundir o leitor incauto.

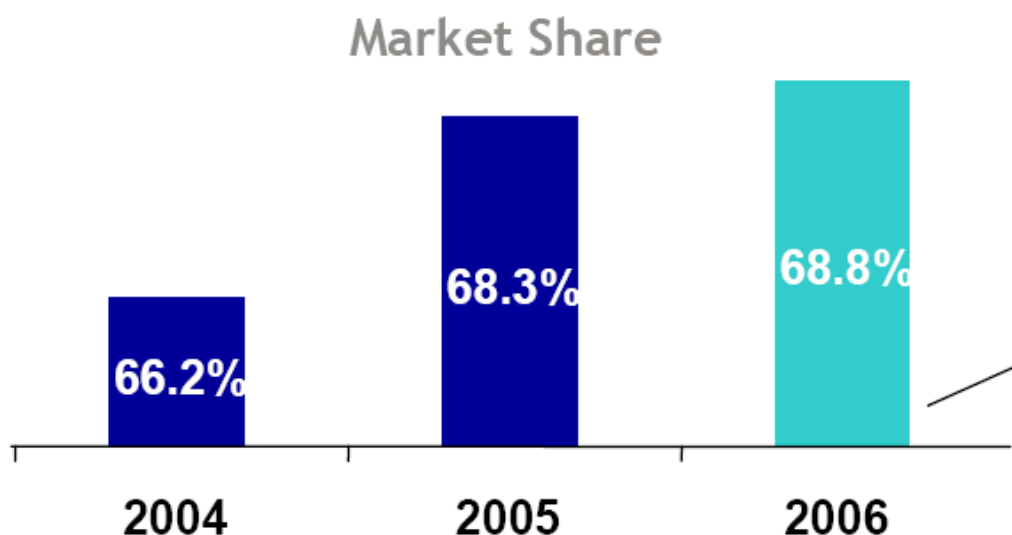



Gráfico da participação de mercado retirado do relatório anual de uma grande empresa brasileira. Um exemplo de gráfico que não começa no zero e faz parecer que as diferenças são maiores do que realmente são.

Gráficos como este são populares em relatórios anuais, pois podem facilmente fazer parecer que o faturamento da empresa aumentou em um terço, quando na realidade subiu apenas um vigésimo.

Dependendo do ponto que você escolhe como base, eles podem exagerar mais ou menos as variações, ou se concentrar em partes especialmente selecionadas da escala, a fim de corroborar o ponto de vista do produtor.

Embora eles sejam tecnicamente corretos, dão margem a conclusões contraditórias e a suspeitas sobre as intenções de quem os desenhou.

Caso 2

Outro exemplo é o gráfico abaixo retirado do relatório anual da mesma empresa do exemplo anterior. Nele são mostrados os custos dos produtos vendidos em Reais, por determinada unidade do produto.

Novamente, a base diferente de zero passa a impressão errada quando comparamos e proporcionamos as barras do gráfico. As variações são visualmente exageradas quando comparadas com as variações dos números.

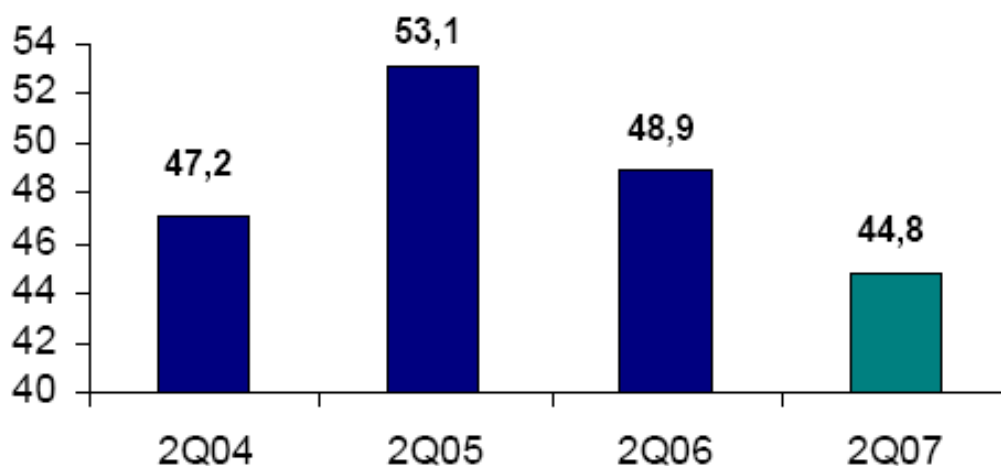



Gráfico com os custos dos produtos vendidos em R$ por determinada unidade do produto, com escala diferente de zero. Mais um exemplo de gráfico enganoso em relatórios anuais.

Este tipo de desenho enganoso é o mais comum. É largamente usado pelos produtores, pois é simples e pode facilmente ser justificado como inocente. Mas, ao mesmo tempo, é facilmente detectado pelo leitor atento.

Caso 3

Outro tipo de gráfico para o qual devemos estar atentos são aqueles que apresentam duas escalas, ilustrando duas variáveis diferentes. Quando estas duas variáveis podem ser comparadas, como quando denotam a mesma grandeza ou unidade de medida, naturalmente tendemos a comparar as curvas ou barras; mas se elas estão em escalas diferentes, essa comparação será enganosa.

O gráfico abaixo mostra a evolução da quantidade de lixo reciclado por uma grande empresa brasileira, e publicado em seu relatório social anual. O eixo da esquerda e as barras ilustram a quantidade de lixo reciclado em toneladas, e o eixo da direita ilustra a porcentagem do total do lixo que é reciclado; ambas as escalas são de base não zero. Isso faz parecer que a quantidade de lixo reciclado aumentou em 400% entre 2004 e 2006, quando na verdade aumentou de menos de 7.000 toneladas para menos de 10.000, um aumento de menos da metade. A outra escala faz parecer que a porcentagem reciclada aumentou em 1/3 do valor anterior, quando na verdade aumentou 1/16.

O problema quando se colocam essas duas medidas relacionadas entre si no mesmo gráfico, em escalas de base não zero, é que as duas escalas estão distorcidas em proporções diferentes. Assim a desinformação se estabelece tanto na comparação entre os anos da mesma variável quanto entre as duas variáveis. Portanto, escalas mistas é outro item com que se deve tomar cuidado em gráficos e diagramas. A não ser que você seja muito vivo em estatística para notar que existem duas escalas diferentes para o mesmo espaço gráfico, poderá ser enganado.

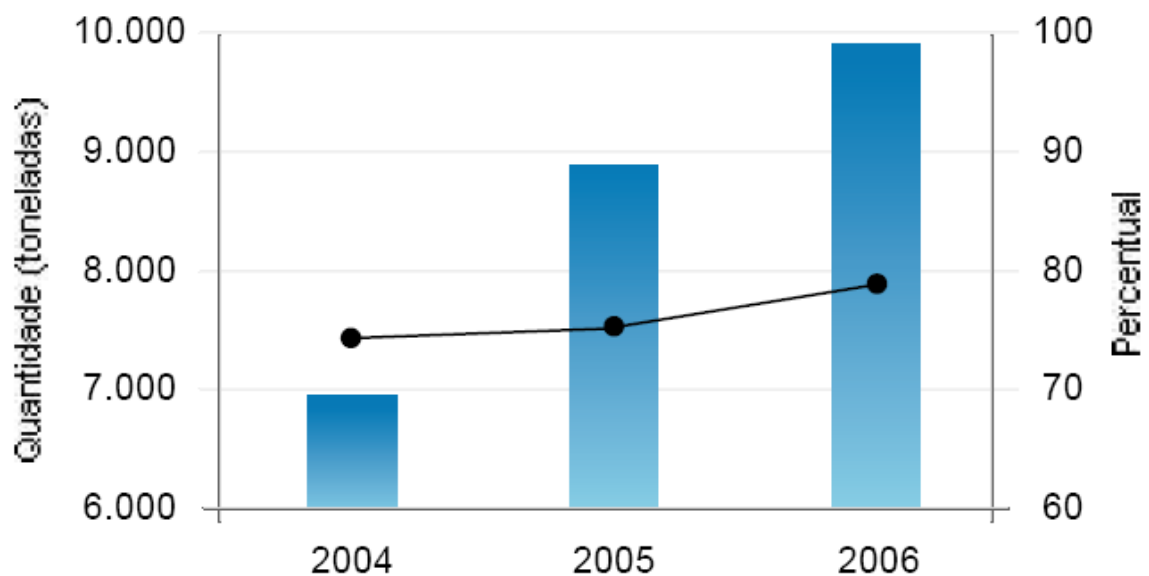



Gráfico do relatório social anual de uma grande empresa brasileira, mostrando a evolução da reciclagem do seu lixo. A escala da esquerda é da quantidade de lixo reciclado, e a escala da direita mostra a porcentagem em relação ao total de lixo. Ambas as escalas são de base não zero.

Caso 4

O gráfico a seguir, do relatório anual de uma grande empresa de entretenimento dos EUA, mostra a evolução das despesas de capital de um determinado departamento. Cada barra representa o valor das despesas de dois setores, e por isso elas são divididas em duas partes.

As barras são circundadas por uma moldura, com espessura considerável. Então o leitor se pergunta: quem fez o gráfico espera que eu leia o valor da barra incluindo a espessura da moldura, ou sem ela? De acordo com a escala, a espessura da moldura é equivalente a US$ 50 milhões, ou pouco mais de 4% do limite da escala. Uma quantia considerável para investidores e analistas de mercado. Mais um exemplo de "lixo de gráfico".

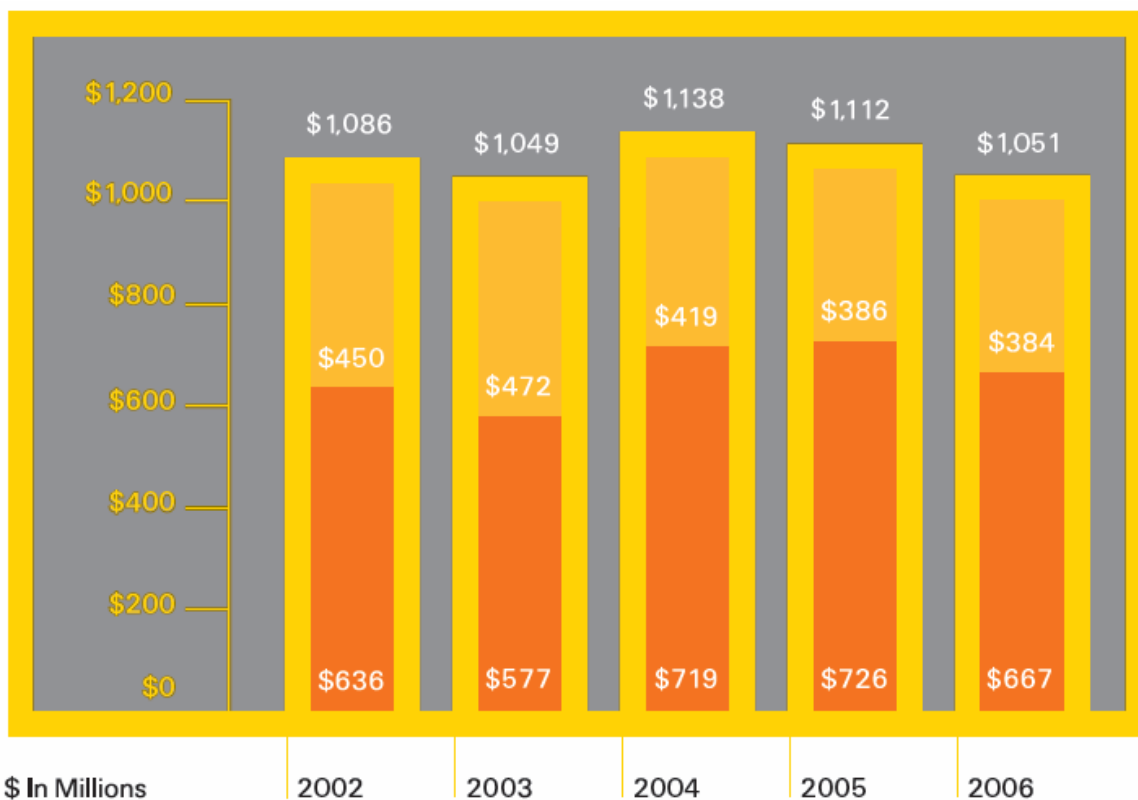



Gráfico das despesas de capital de um setor de uma grande empresa de entretenimento dos EUA. O uso de uma moldura em torno das barras confunde o leitor.

Caso 5

O gráfico de barras a seguir apresenta a evolução do endividamento bruto em bilhões de Reais de uma grande estatal brasileira. As barras são divididas em três partes, mostrando parcelas de tipos diferentes de endividamento e uma linha no alto de cada barra mostra o valor total para aquele ano. Cada parte de cada barra é separada da seguinte por um espaço, que tem o mesmo tamanho em todas as vezes que aparece.

O problema desse gráfico com barras flutuantes é que os espaços vazios aumentam a altura total da barra e inviabilizam a comparação direta entre cada componente da barra e o total, passando a impressão de que cada parte é menor do que realmente é em comparação ao todo. Isto pode ser conveniente num gráfico de endividamento que mostra uma tendência de aumento.

As barras flutuantes impedem o leitor de comparar os valores numéricos pela simples comparação do comprimento de cada barra, que é justamente uma das principais vantagens desse tipo de gráfico.

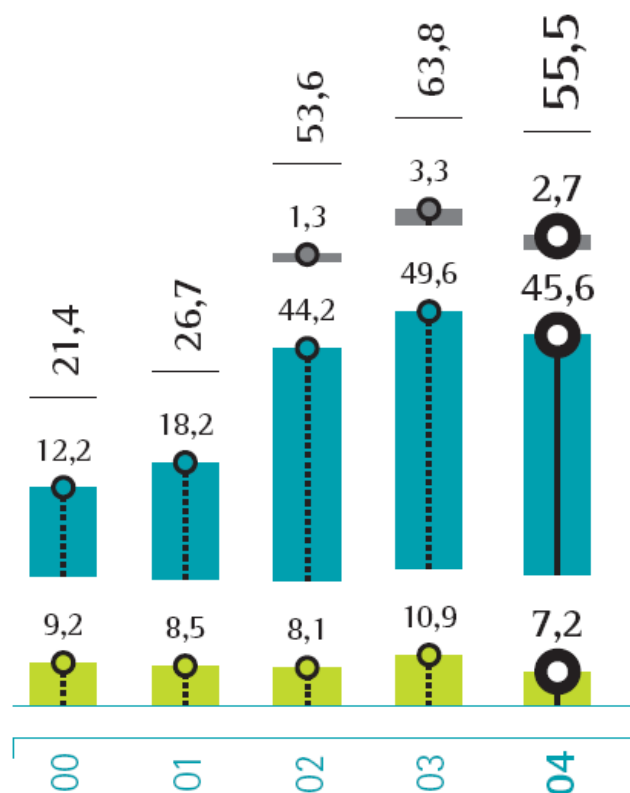



Gráfico da evolução do endividamento bruto em bilhões de reais, retirado do relatório anual de uma grande estatal brasileira. O uso de barras flutuantes impede a comparação entre as partes e o total.

Caso 6

O próximo gráfico foi tirado do relatório anual de uma tradicional empresa automobilística dos EUA. O interessante deste gráfico é que a evolução dos ganhos por ação é mostrada com o ano mais recente à esquerda e os anos anteriores à direita, uma disposição inversa à nossa noção de que a direção de leitura, ou sequência, se realiza da esquerda para a direita, ou mesmo da nossa intuição de que o tempo deve ser representado passando da esquerda para a direita. Essa disposição é uma escolha bastante suspeita, considerando que a empresa não tem números bons para mostrar, e este gráfico pode fazer parecer que o grande número negativo à esquerda estava no passado, e que hoje a empresa voltou a ter resultados positivos, o que na realidade é exatamente o contrário.

Gráfico dos ganhos por ação do relatório anual de uma grande empresa automobilística dos EUA. Note que a evolução dos ganhos é mostrada com os anos decrescendo da esquerda para a direita, fazendo parecer que os ganhos estão subindo, enquanto na verdade se tornaram bem negativos no último ano.

Caso 7

O gráfico seguinte é da mesma empresa do exemplo anterior, mas de um relatório anual anterior. Além de apresentar o mesmo problema que foi descrito anteriormente, este gráfico apresenta outro. O gráfico mostra os valores dos ganhos anuais por ação de três anos consecutivos, representados por linhas, que juntas formam um desenho que lembra uma autoestrada; a linha do meio é tracejada, lembrando a faixa que permite ultrapassagens nas estradas.

Este é um exemplo de como a preocupação com o estilo pode superar a preocupação com a compreensão. Embora realmente seja uma ideia muito criativa fazer um gráfico em forma de autoestrada no relatório anual de uma empresa automobilística, o problema é que a linha tracejada dá um peso visual menor para o valor da linha do meio, diminuindo visualmente a sua importância relativa às outras linhas. Novamente uma escolha muito suspeita, considerando-se que o ano representado no centro foi o que teve o pior desempenho.

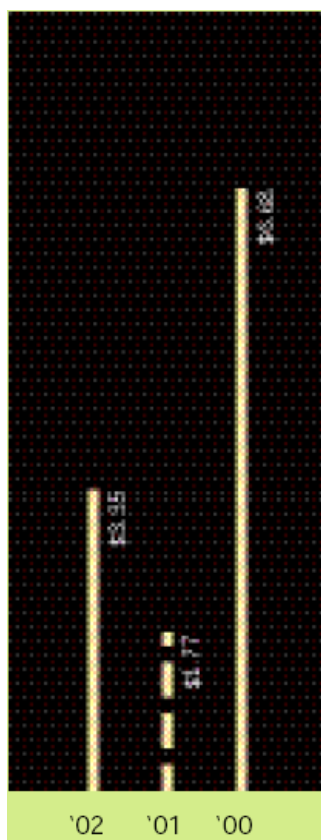



Gráfico dos ganhos anuais por ação de uma grande empresa automobilística dos EUA, cujas linhas que representam o valor de três anos consecutivos, formam o desenho de uma autoestrada. Um exemplo de preocupação com o estilo ao invés da compreensão.

Caso 8

O gráfico a seguir é um caso onde não há erros, mas que poderia ser facilmente melhorado. Ele mostra o valor dos estoques de uma grande estatal brasileira, em gráfico de torta, com apenas quatro valores. Ele poderia ser facilmente rotulado diretamente, dispensando a legenda logo em baixo, facilitando a sua leitura e comparação entre as partes.

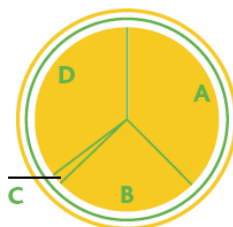
Estoques – consolidado – 31/12/2003
R$ Milhões

A) Matéria-prima – **3.919**
B) Derivados – **2.607**
C) Álcool – **195**
D) Outros – **3.674**

Gráfico de torta retirado do relatório anual de uma estatal brasileira. Ele poderia ser melhorado com uma rotulação direta, dispensando a legenda logo abaixo.

Caso 9

Já o gráfico abaixo, que mostra os benefícios a empregados de uma grande empresa brasileira, apresenta a rotulação direta ao invés da legenda, o que facilita a leitura. No entanto, há muitas fatias pequenas (entre elas há quatro fatias com o valor de 1%) que são difíceis de visualizar, o que dificulta a comparação entre cada fatia e o todo.

Este não é um caso extremo, como o dos gráficos de torta com dezenas de fatias; mesmo este tipo de gráfico ainda se mostra adequado para esta situação. Entretanto, o uso de uma tabela, ou um gráfico de barras, seriam outras boas opções.

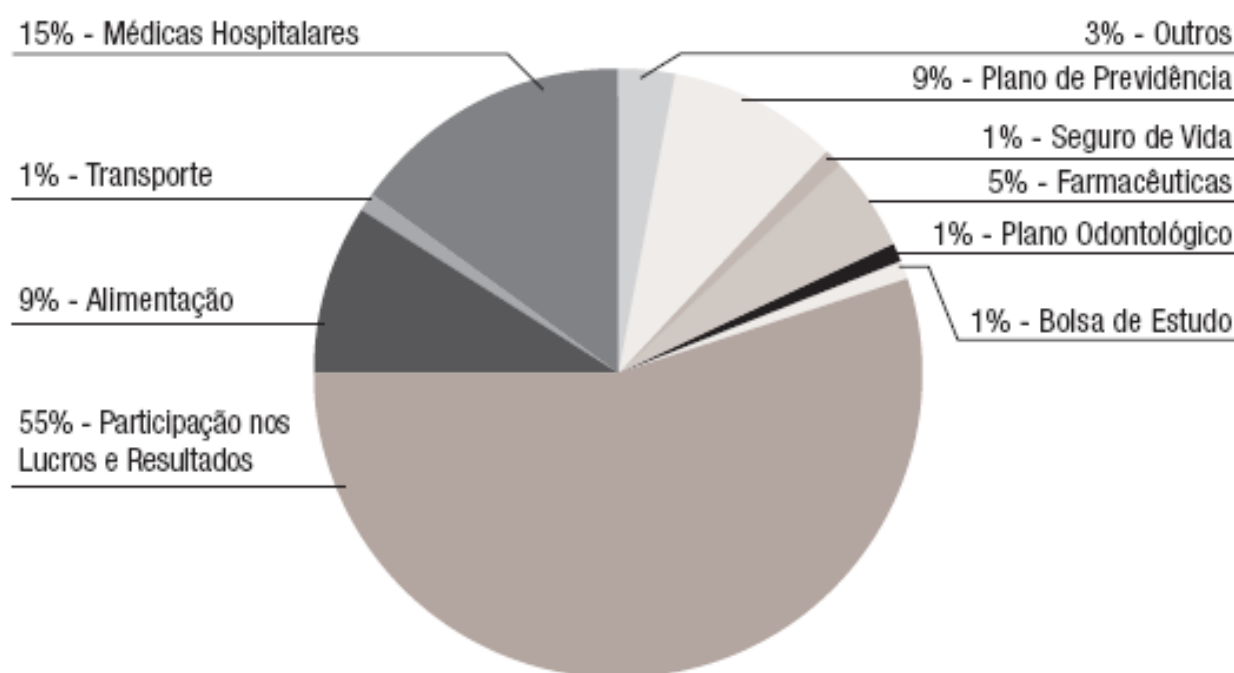



Gráfico de barras do relatório social anual de uma grande empresa, mostrando os benefícios a empregados. Note a quantidade de fatias muito pequenas, que dificultam a comparação entre as partes e o todo.

Caso 10

O gráfico a seguir é um exemplo de uso equivocado dos gráficos de tortas. Este tipo de desenho deve ser usado para representar proporções, relações, porcentagens, ou qualquer outra medida de relação entre um todo e suas partes.

Este gráfico não indica a relação entre as fatias e o todo. O leitor só sabe o valor absoluto de cada fatia. Para saber o valor absoluto de toda a torta, deverá somar o valor de todas as fatias, e caso deseje saber a relação entre as partes e o todo, deverá dividir o valor de cada fatia pelo valor total da torta. Depois de constatado todo este trabalho, não seria melhor ter apresentado os valores um uma tabela? Afinal, o objetivo deste tipo de gráfico, que é a comparação, foi totalmente derrubado pela inabilidade do produtor.

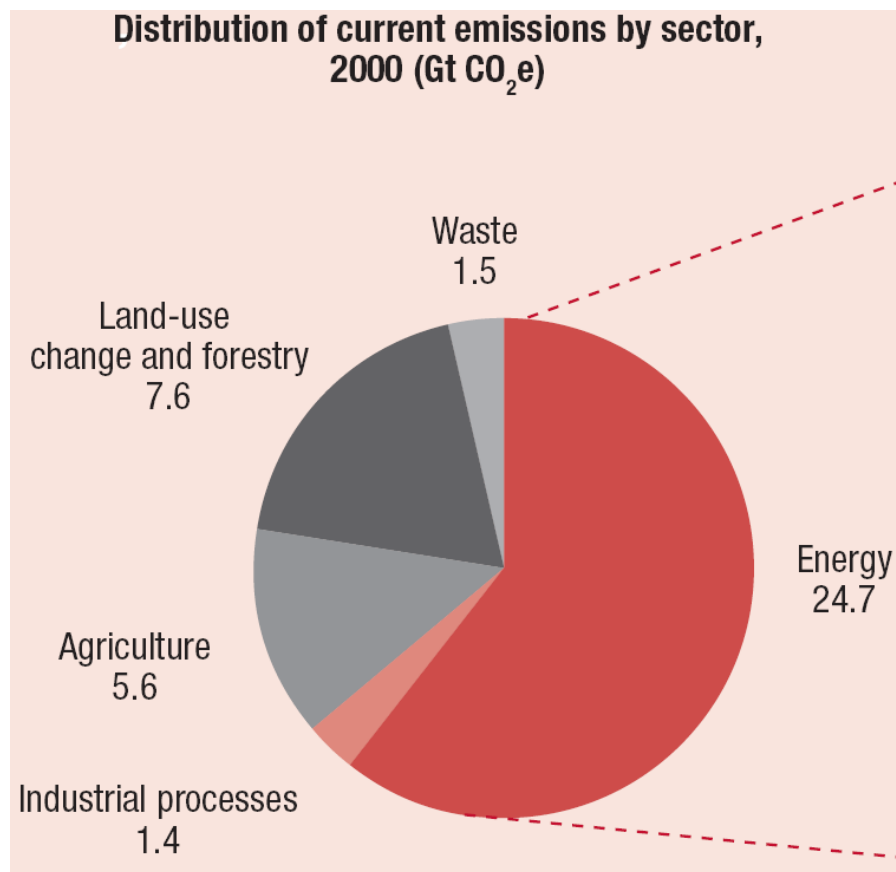



Gráfico retirado do relatório de um programa das Nações Unidas. Gráficos de tortas podem até indicar valores absolutos, desde que indiquem as relações em primeiro plano, para dar o enfoque correto ao leitor.

Caso 11

O gráfico de torta abaixo foi retirado do relatório social anual de uma grande empresa brasileira e mostra a relação dos valores distribuídos aos segmentos da sociedade. Este é um claro exemplo de uso desnecessário de uma perspectiva em 3D, que não adiciona nenhuma informação, somente "lixo de gráfico".

Este gráfico sofre de "despenhadeirite", que é esta representação de altura que aparece em torno da torta, mas que não traz nenhum valor numérico e só confunde e distrai.

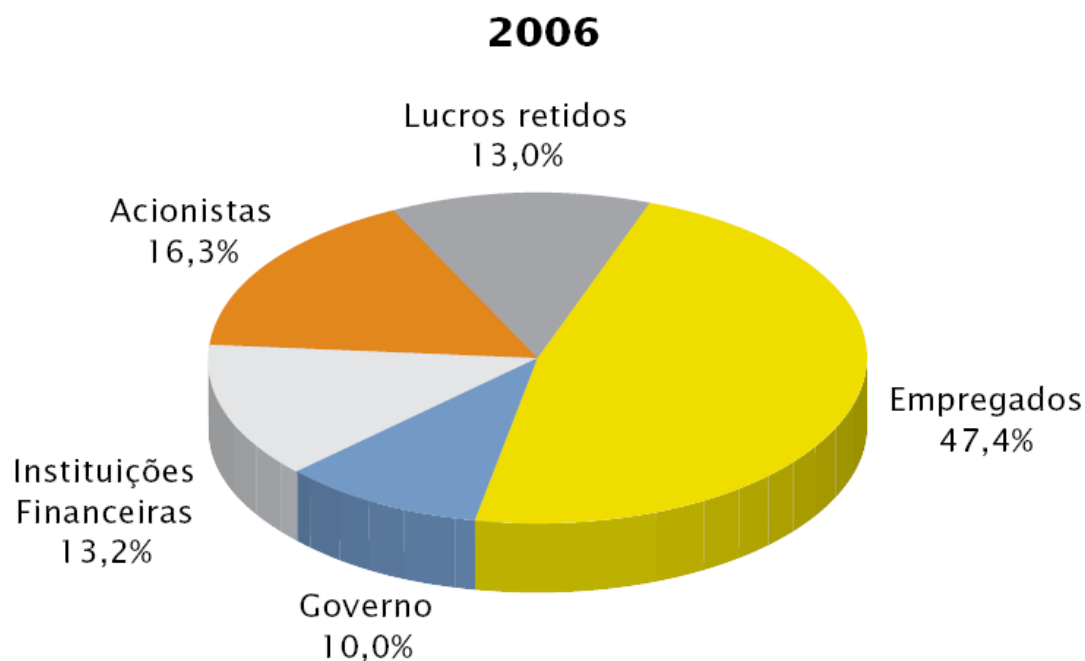



Gráfico retirado do relatório social anual de uma grande empresa, ilustrando a distribuição dos lucros nos segmentos da sociedade. Um exemplo de uso desnecessário de perspectiva em 3D.

Além disso, a representação em perspectiva dos gráficos em 3D distorce o tamanho relativo das partes do gráfico, conferindo maior tamanho às fatias que estão na frente, e relativamente diminuindo as fatias na parte de trás. As pessoas vão pensar que as fatias de baixo são as mais importantes.

No gráfico anterior, as fatias enfatizadas na parte da frente são os lucros repassados ao governo e aos funcionários, e a fatia que está na parte de trás é a dos lucros retidos pela empresa, uma coincidência bastante suspeita.

É um truque comum colocar a fatia que você quer enfatizar na parte de baixo da torta, e aumentar o quanto for possível a altura da torta, e incliná-la o máximo para trás, como no gráfico a seguir. Nele, a distorção das fatias é bastante evidente.

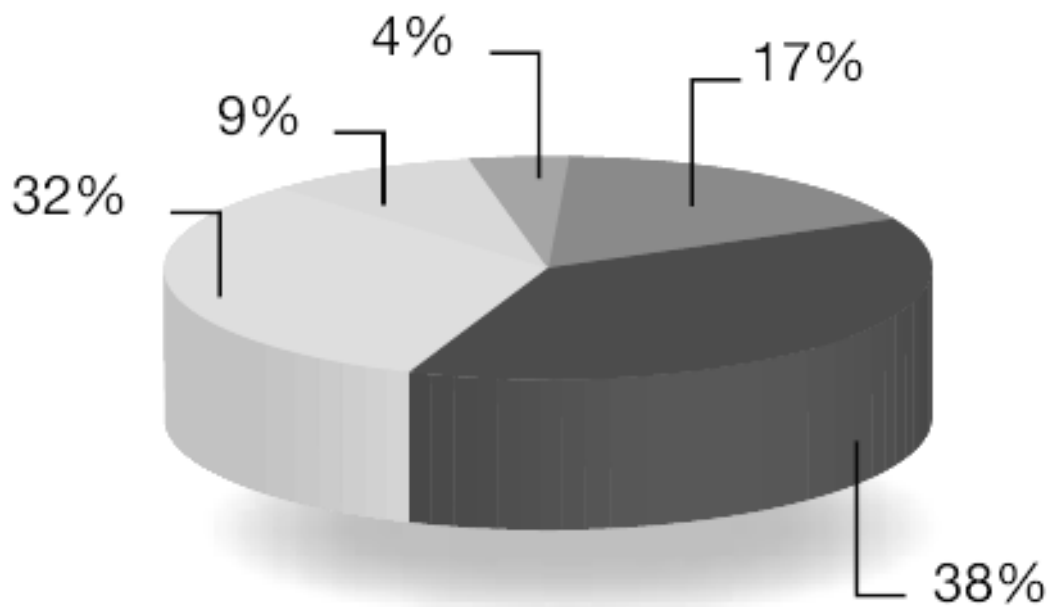



Gráfico retirado do relatório de um órgão governamental europeu. As distorções pela perspectiva são aumentadas com o aumento da altura e da inclinação da torta. Note que o produtor ainda teve o cuidado de adicionar a sombra da torta. Ao confeccionar este gráfico, acredito, já estava sofrendo de insolação.

Caso o produtor ainda não esteja satisfeito com o grau de distorção causada pela perspectiva num gráfico de torta, ele pode colocá-la na vertical e transformar a torta numa roda. As fatias da parte de baixo da roda são apresentadas com um tamanho relativo menor em comparação com as fatias da parte de cima.

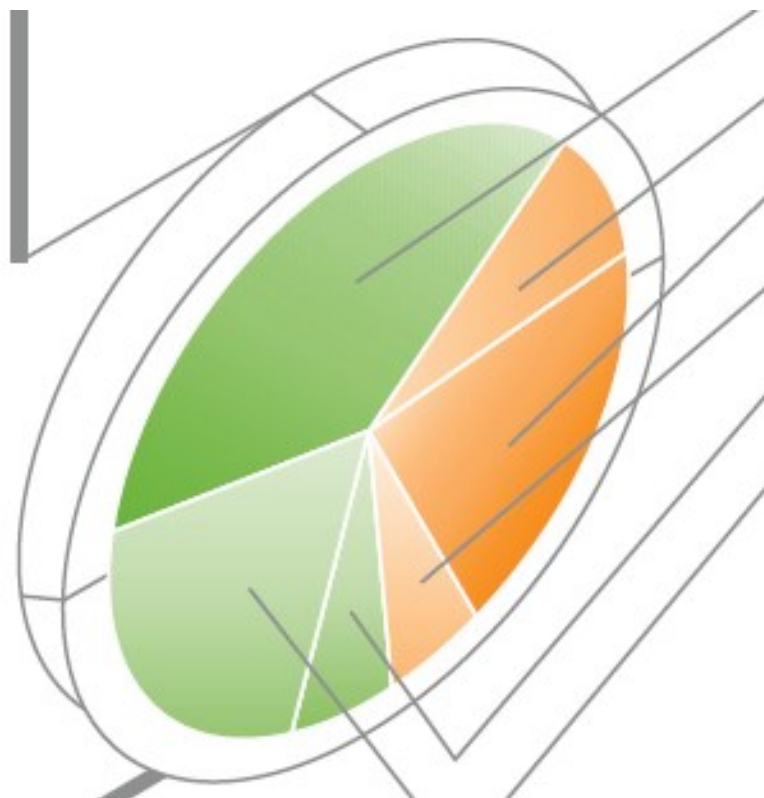



Gráfico retirado do relatório anual de uma grande estatal brasileira. Colocar as tortas na vertical e transformá-las em roda agrava ainda mais os problemas do leitor, que além de continuar sendo enganado pela perspectiva, não está

familiarizado com este desenho não usual.

Mesmo com todos os efeitos tridimensionais que os computadores nos oferecem ao esforço de um *click*, parece que essas opções não são mais suficientes para saciar a sede dos produtores que estão mais interessados em valorizar o recipiente do que o conteúdo, ao preço de adicionar mais "lixo de gráfico". O produtor do gráfico a seguir não ficou satisfeito com a perspectiva em 3D do gráfico de torta, e decidiu adicionar imagens de fotos como textura para cada fatia.

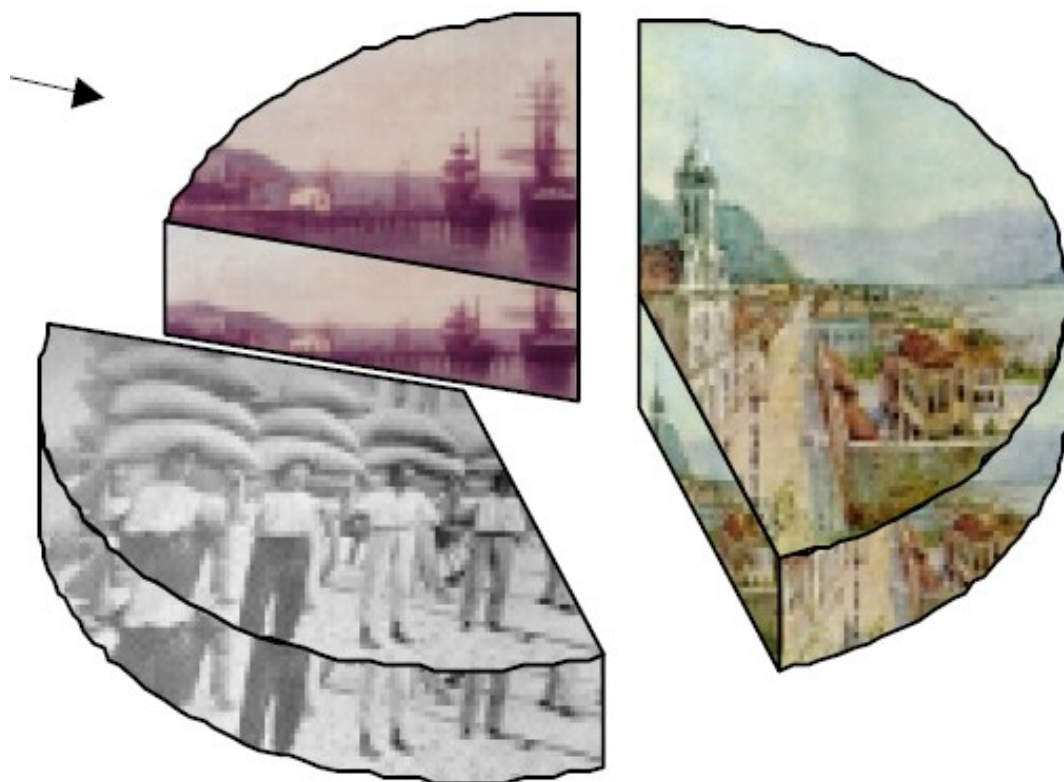



Gráfico retirado do relatório anual da administradora de um porto brasileiro. As fatias ao invés de serem coloridas, são cobertas pela imagem de fotos. Mais um exemplo de preocupação com o estilo que supera a preocupação com a compreensão.

Não é porque os recursos computacionais disponíveis nos permitem criar tais "obras de arte" que devemos considerar que elas são as melhores opções para levarem ao entendimento, e nem devemos cair na tentação de usá-las nos casos em que não agregam informação. Gráficos mais bonitos não são intrinsecamente melhores. Quase todas as demandas por gráficos são satisfeitas por desenhos simples em duas dimensões, que não são tão chamativos, mas levam o leitor à compreensão.

Caso 12

O gráfico de rosca a seguir mostra a distribuição das vendas entre as regiões brasileiras de uma estatal brasileira. É um exemplo de gráfico que poderia ser perfeitamente projetado em uma perspectiva em 2D, mas a preocupação com o estilo e de fazer o relatório anual parecer legal e moderno, fez com que se escolhesse a representação em 3D que somente distorce o gráfico.

O gráfico está distorcido, mas de forma irregular. Não é possível concluir se o gráfico está inclinado para a esquerda, direita, ou se está em forma oval. A espessura da rosca parece variar também. Este gráfico é quase um desenho projetado para causar ilusão de ótica. A fatia da região Sudeste que é indicada com 56,2%, no gráfico parece contar com menos da metade da área, enquanto os 5,4% da região Norte parecem maior que o devido.

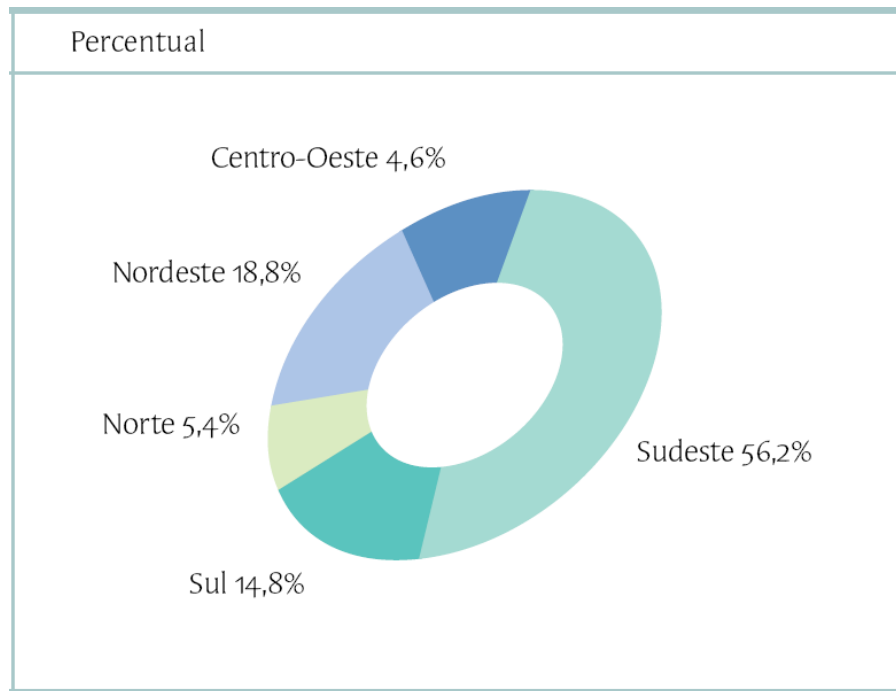



Gráfico do relatório anual de uma estatal brasileira, mostrando a distribuição das vendas entre as regiões do Brasil. Outro exemplo de como o uso desnecessário da perspectiva em 3D deforma os gráficos.

Caso 13

A ânsia por criar gráficos em três dimensões é tão grande que não se contenta com gráficos circulares transformados em tortas flutuantes ou em rodas. É também necessário dar profundidade à linha de um gráfico de evolução, que teoricamente não tem profundidade.

O gráfico abaixo transforma uma linha num plano quando adiciona a dimensão de profundidade, o que não traz informação nova, mas apenas confusão. É esperado do leitor que ele acompanhe a linha formada pela parte da frente ou de trás do plano?

Além disso, a escala vertical não começa no zero, aumentando a percepção da variação. Se colocada numa escala de base zero, o gráfico mostraria quase uma constante.

Este gráfico apresenta apenas seis números (três anos e três valores) e ocupa uma área considerável do relatório. É um exemplo de representação de dados de baixíssima densidade. Poucas publicações têm uma densidade de dados por centímetro quadrado tão baixo como este gráfico. Seria muito melhor ter usado uma tabela, que ocuparia uma fração do espaço e realizaria o mesmo trabalho.

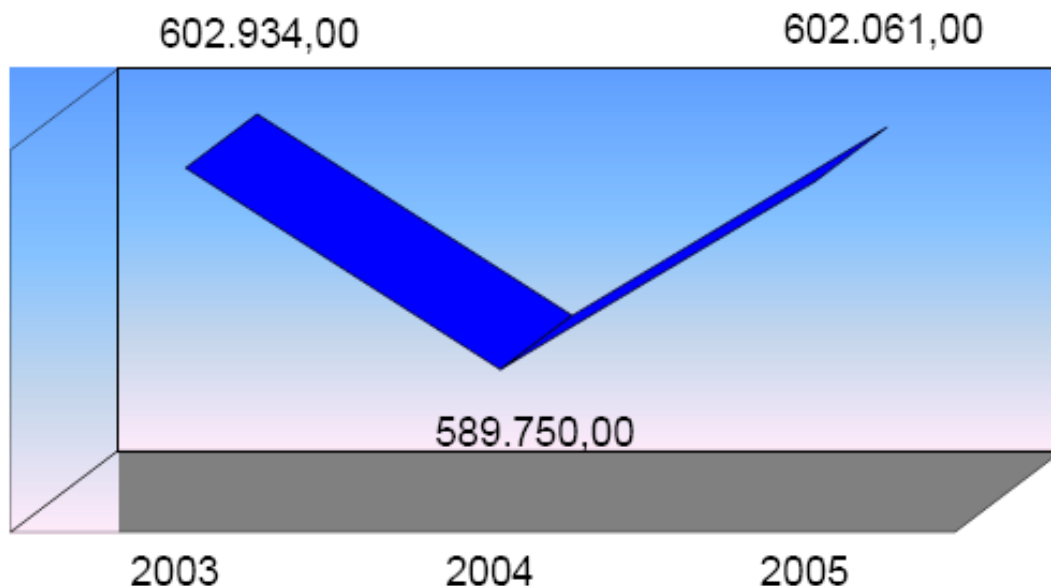



Gráfico do relatório de um programa social do governo brasileiro mostrando a evolução dos recursos aplicados. Um gráfico de evolução espaçoso e em 3D para representar apenas seis números. Uma tabela teria realizado o mesmo trabalho ocupando apenas uma fração do espaço deste gráfico.

Caso 14

Geralmente não há uma regra para escolher a perspectiva em gráficos 3D. Um truque usado em gráficos de barras 3D é escolher uma perspectiva que faça a face de trás da barra estar num nível mais alto que a face da frente. Então o observador não sabe se o que deve ser comparado com a escala é a face da frente ou a de trás, que é mais alta e por isso representa um valor maior. Outros truques de perspectiva podem esconder as faces laterais ou criar outras relações entre a face frontal e traseira.

O gráfico a seguir apresenta dois problemas. O primeiro problema é que a escala não começa no zero, um assunto que já foi discutido. O outro problema é que a face de trás das barras é mais alta que a face da frente. O produtor do gráfico espera que eu leia o valor da barra pela altura de qual das faces? Para a barra mais à esquerda, a face frontal indica 17,8, e a de trás indica 18. Pelo menos o produtor teve o cuidado de colocar rotulações diretas sobre cada barra, indicando os seus valores. Este recurso ameniza o problema da perspectiva errada, no entanto não elimina o problema das barras inutilmente representadas em 3D.

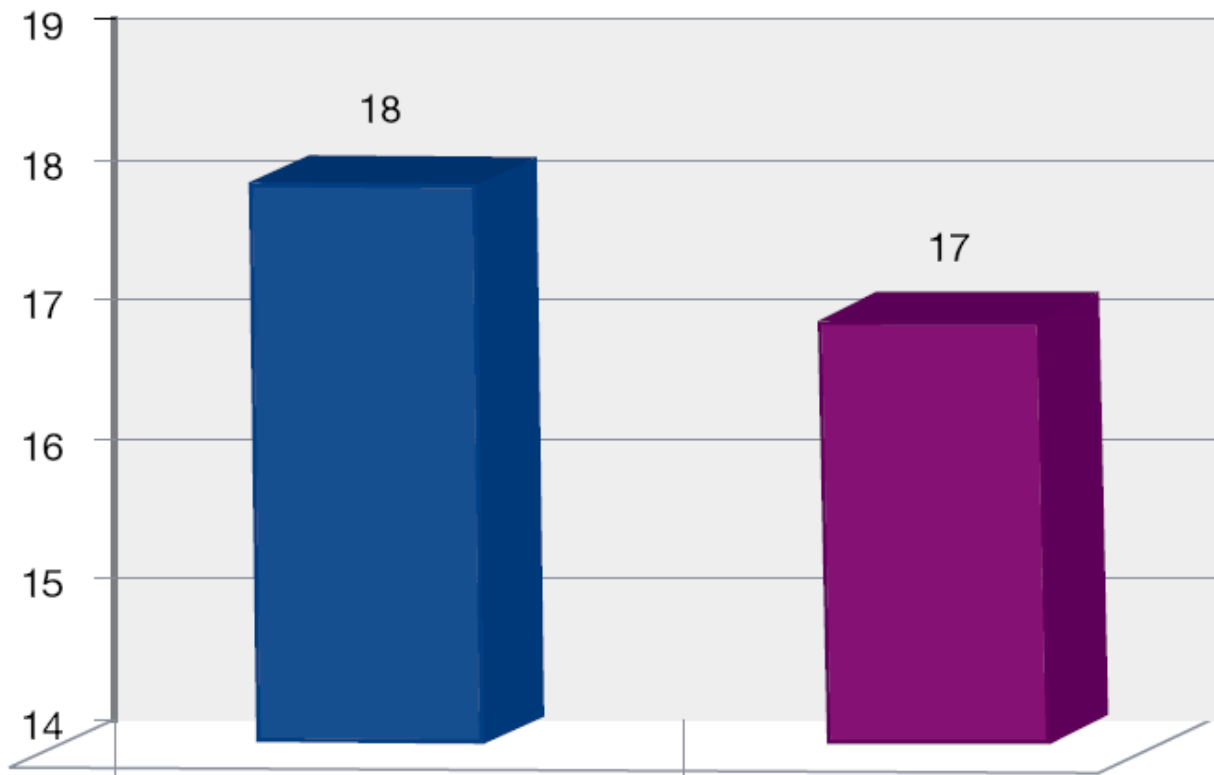



Gráfico do relatório de um órgão governamental europeu, com barras tridimensionais cuja perspectiva deixa a face de trás mais alta que a da frente. Qual das faces indica o valor correto?

Se o uso de um gráfico de barras em 3D for realmente imprescindível, certifique-se de que a perspectiva deixe a face da frente mais alta, para não enfatizar a face traseira e confundir o observador quanto ao valor. Se necessário indique o valor de cada barra com um número no topo, como neste caso. Mas a melhor escolha continua sendo um gráfico em 2D que economiza todo esse trabalho.

Infelizmente as barras em 3D estão cada vez mais populares nos ambientes corporativos e gerentes estão realmente tomando decisões operacionais e estratégicas baseados neste tipo de gráfico. É assustador.

Caso 15

Caso você ainda não esteja satisfeito com a aparência do seu gráfico de barras em 3D, você pode adicionar texturas de tijolos às barras, e fazê-las parecerem torres ou colunas que foram erguidas tijolo a tijolo. Talvez o produtor quisesse fazer uma metáfora do esforço necessário para alcançar tais resultados com o esforço de erguer tal construção.

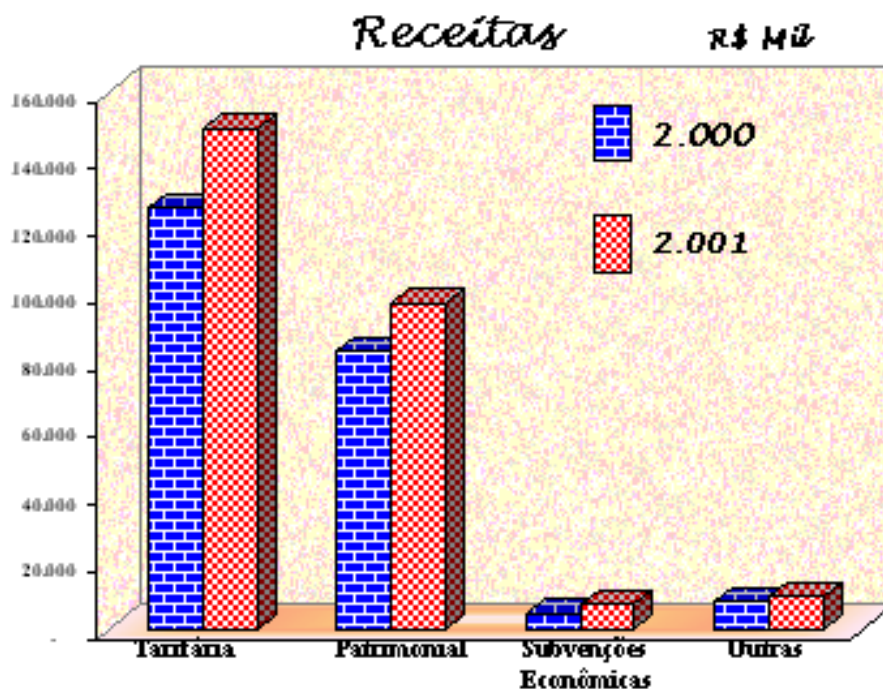



Gráfico retirado do relatório anual da administradora de um porto brasileiro. Texturas foram adicionadas às barras para dar a aparência de sólidas construções. Mais um exemplo de "lixo de gráfico".

Essas texturas, emaranhados de linhas e pontos, confundem e fazem doer os olhos do observador, além de poluírem visualmente o gráfico. São elementos gráficos que não adicionam informação nova, e por isso são mais um exemplo de "lixo de gráfico".

Caso 16

O gráfico de barras a seguir, retirado do relatório anual de uma grande estatal brasileira, mostra a evolução do volume de produção. Novamente o uso desnecessário da perspectiva em 3D distorce o gráfico, lançando suspeitas sobre as intenções do produtor. Só porque os computadores têm capacidade para produzir tais gráficos, não significa que devemos testar os seus limites e produzir gráficos tão elaborados quando um gráfico tradicional iria ser bem mais adequado.

As barras repousam sobre um eixo que sobe da esquerda para a direita, fazendo com que as barras da direita terminem em alturas relativamente superiores às suas vizinhas da esquerda, passando a impressão para o leitor distraído que houve um aumento de produção maior que o ocorrido. Note como a barra de 2002 termina em uma altura muito superior que a barra de 1999, sugerindo um aumento de cerca de 50%, enquanto na verdade houve um aumento de menos de 30% em relação à produção de 1999. Pelo menos há várias linhas da escala vertical que cortam o plano de fundo do gráfico e ajudam o leitor a manter o rumo enquanto compara as barras, mitigando as distorções que a representação de profundidade está causando. Se a projeção em 2D tivesse sido adotada, essas linhas seriam dispensáveis, diminuindo os elementos gráficos que não trazem informação, ou se preferir, o "lixo de gráfico".

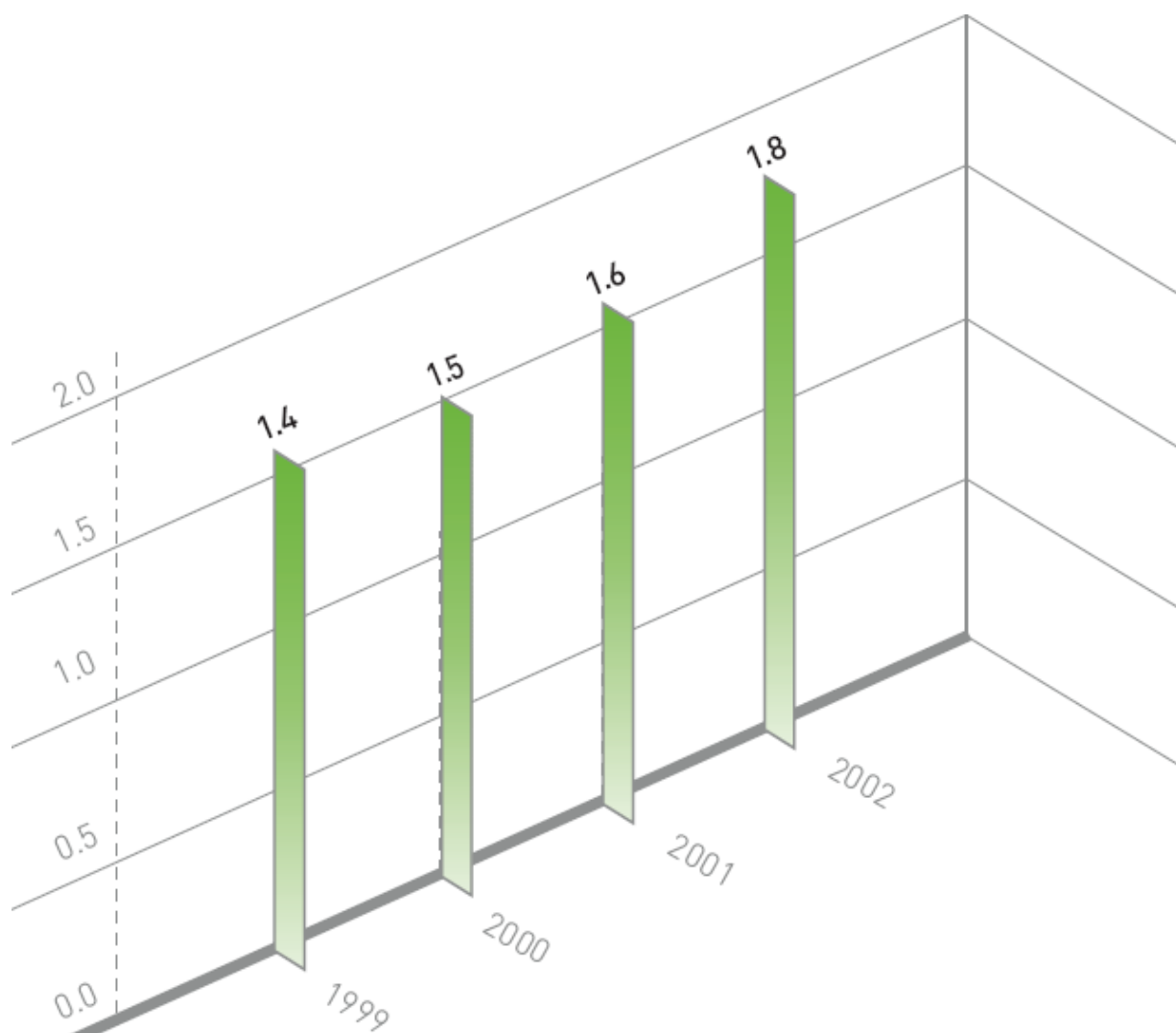



Gráfico do relatório anual de um grande estatal brasileira, mostrando a evolução do volume de produção. O uso desnecessário de uma perspectiva em 3D novamente distorce as relações e engana o leitor. Neste caso poderia sugerir um aumento de produção maior que o indicado pelos números.

Conclusão

São na política, na economia, na publicidade, na imprensa e em muitas outras áreas que estão sempre nos tentando convencer de alguma coisa. E nada melhor para isso do que alguns gráficos e porcentagens cuidadosamente selecionados para corroborarem com os argumentos. Isto é especialmente efetivo nas mídias sociais, um meio muito popular e propício a leituras rápidas e sem crítica. Mas você, leitor, que chegou até o final deste pequeno livro, está muito mais preparado para identificar essas artimanhas e armadilhas.

E não restrinja-se aos gráficos e estatísticas. Este livro também é um convite ao pensamento crítico. Questione as notícias ruins demais, os argumentos bons demais, as propostas generosas demais e as histórias fantásticas demais; na maioria das vezes, uma pesquisa de alguns minutos na Internet é suficiente para clarificar o assunto. Afinal, com a Internet nunca foi tão fácil espalhar a mentira, mas também descobrir a verdade.

Alessandro Nicoli de Mattos

Sobre o Conteúdo Não Original

Notas sobre as condições de utilização dos conteúdos que não são de propriedade do autor desta obra:

Fotos e figuras: todas as fotos e figuras usadas são de domínio público ou de divulgação, e suas fontes estão indicadas nas notas.

Gráficos analisados: os gráficos analisados em Sobre Estatísticas, Gráficos e Mentirosos são retirados de relatórios anuais de empresas e relatórios do governo disponíveis na Internet. Suas fontes não são citadas, assim como não procurei autorização para usá-las por motivos óbvios, uma vez que eles são analisados pela (falta de) integridade e qualidade.

Notas

[01] Instituto Paulo Montenegro, Indicador de Alfabetismo Funcional. Disponível em *http://www.ipm.org.br*.

[02] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 47.

[03] HUFF, Darrel, *How to Lie with Statistics*, 1954. Veja mais informações sobre este livro na bibliografia.

[04] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 8.

[05] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 13.

[06] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 28.

[07] PLATT, Washington. *A Produção de Informações Estratégicas*. Rio de Janeiro: Agir Editora, 1974. Figura da pág. 225.

[08] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 42 e 43.

[09] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 122-142.

[10] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 123.

[11] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 131.

[12] TUFTE, Edward Rolf, *Visual and Statistical Thinking; Displays of Evidence for Making Decisions*. Cheshire: Graphics Press, 2005, pág. 23.

[13] HUFF, Darrel, *How to Lie with Statistics*. New York: W W Norton & Company INC, 1954, pág. 60.

[14] Instituto Paulo Montenegro, Indicador de Alfabetismo Funcional. Disponível em *http://www.ipm.org.br*.

[15] TUFTE, Edward Rolf, *The Visual Display of Quantitative Information*. Cheshire: Graphics Press, 2007, pág. 53.

[16] TUFTE, Edward Rolf, *The Visual Display of Quantitative Information*. Cheshire: Graphics Press, 2007, pág. 51.

[17] *O Estado de São Paulo*, 16 de setembro de 2007, pág. B4. Imagem digitalizada a partir do original.

[18] *O Estado de São Paulo* 16 de setembro de 2007, pág. B20. Imagem digitalizada a partir do original.

[19] *Guia de Investimentos em Ações – BB Estilo*. Imagem digitalizada a partir do original.

[20] JONES, Gerald Everett. *How to Lie with Charts: Second Edition: Second Edition*. Santa Monica: La Puerta Productions, 2007, pág. 76 e 77.

[21] WURMAN, Richard Saul, DIVERSOS, 2000. *Understanding USA*. Excerto da pág. 4 do capítulo 2. Disponível em *http://www.understandingusa.com*.

[22] *Annuario de Estatística Demographo-Sanitaria de 1908*, pelo Dr. Cássio de Rezende, Imprensa Nacional, Rio de Janeiro, 1910. Figura entre as páginas 112 e 113. Documento disponível em *http://memoria.nemesis.org.br*.

[23] *Annuario de Estatística Demographo-Sanitaria de 1915-1916*, pelo Dr. Sampaio Vianna, Imprensa Nacional, Rio de Janeiro, 1926. Figura entre as páginas 114 e 115. Documento disponível em *http://memoria.nemesis.org.br*.

[24] *O Estado de São Paulo*, 9 de março de 2008, pág. C1. Imagem digitalizada a partir do original.

[25] TUFTE, Edward Rolf, *The Visual Display of Quantitative Information*. Cheshire: Graphics Press, 2007, pág. 71.

[26] JONES, Gerald Everett. *How to Lie with Charts: Second Edition: Second Edition*. Santa Monica: La Puerta Productions, 2007, pág. 19.

[27] *O Estado de São Paulo*, 16 de setembro de 2007, pág. Ce4. Imagem digitalizada a partir do original.

[28] *Annuario de Estatística Demographo-Sanitaria de 1908*, pelo Dr. Cássio de Rezende, Imprensa Nacional, Rio de Janeiro, 1910. Figura entre as páginas 26 e 27. Documento disponível em *http://memoria.nemesis.org.br*

[29] *O Estado de São Paulo*, 16 de setembro de 2007, pág. B4. Imagem digitalizada a partir do original.

[30] *O Estado de São Paulo*, 16 de setembro de 2007, pág. B20. Imagem digitalizada a partir do original.

[31] *Valor Econômico*, 25 de setembro de 2007, pág. C9. Imagem digitalizada a partir do original.

[32] TUFTE, Edward Rolf, *Visual and Statistical Thinking; Displays of Evidence for Making Decisions*. Cheshire: Graphics Press, 2005, pág. 27.

[33] JONES, Gerald Everett. *How to Lie with Charts: Second Edition: Second Edition*. Santa Monica: La Puerta Productions, 2007,

pág. 43 e 44.

[34] JONES, Gerald Everett. *How to Lie with Charts: Second Edition: Second Edition*. Santa Monica: La Puerta Productions, 2007, pág. 44.

[35] *O Estado de São Paulo*, 23 de março de 2008, pág. B4. Imagem digitalizada a partir do original.

[36] JONES, Gerald Everett. *How to Lie with Charts: Second Edition*. Santa Monica: La Puerta Productions, 2007, pág. 40.

[37] *Annuario de Estatística Demographo-Sanitaria de 1915-1916*, pelo Dr. Sampaio Vianna, Imprensa Nacional, Rio de Janeiro, 1926. Figura entre as páginas 30 e 31. Documento disponível em *http://memoria.nemesis.org.br*.

[38] TUFTE, Edward Rolf, *The Visual Display of Quantitative Information*. Cheshire: Graphics Press, 2007, pág. 68.

[39] Gráfico disponível em *http://www.planejamento.sp.gov.br/PlanOrca/orca.asp*.

[40] WURMAN, Richard Saul. *Ansiedade de Informação: Como transformar informação em compreensão*. São Paulo: Cultura Editores Associados, 1991, pág. 301.

[41] JONES, Gerald Everett. *How to Lie with Charts: Second Edition*. Santa Monica: La Puerta Productions, 2007, pág. XVII.

[42] Perozzo, Luigi. *Statistica Grafica - Della rappresentazione grafica di una collettività di individui nella successione del tempo, e in particolare dei diagrammi a tre coordinate - Memoria di Luigi Perozzo. Annali di Statistica*, Serie 2, Vol. 12, 1880 (*Ministero d'Agricoltura, Industria e Commercio, Direzione di Statistica*). Imagem disponível em *http://www.math.yorku.ca/SCS/Gallery/images/stereo2.jpg*.